7/973

183438

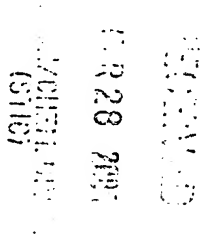
STIC-Biotech/ChemLib

From: Huff, Sheela
Sent: Tuesday, March 28, 2006 10:39 AM
To: STIC-Biotech/ChemLib
Subject: search for 10609217

please search and interference search SEQ ID No. 83-85. 124, 419, 420, 421, 41, 339 and 340.

Thanks

Sheela Huff
Art Unit 1643
571-272-0834
Remsen 3A15
mailbox Remsen 3C18



****
Searcher:_____
Searcher Phone:_____
Date Searcher Picked up:_____
Date completed:_____
Searcher Prep Time:_____
Online Time:_____

****
Type of Search
NA#_____ AA#:_____
S/L:_____ Oligomer:_____
Encode/Transl:_____
Structure #:_____Text:____
Inventor:_____ Litigation:___

****
Vendors and cost where applicable
STN:_____
DIALOG:_____
QUESTEL/ORBIT:_____
LEXIS/NEXIS:_____
SEQUENCE SYSTEM:_____
WWW/Internet:_____
Other (Specify):_____

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: March 31, 2006, 16:09:06 ; Search time 42.4129 Seconds
(without alignments)
113.955 Million cell updates/sec

Title: US-10-609-217-41
Perfect score: 54
Sequence: 1 GVREVIVMHML 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	54	100.0	11	2	AAY22362	Aay22362 TPO recep
2	54	100.0	11	3	AAB16985	Aab16985 TPO-mimet
3	54	100.0	11	5	ABB72871	Abb72871 TPO mimet
4	54	100.0	11	7	ADJ73022	Adj73022 TPO mimet
5	54	100.0	11	8	ADJ52657	Adj52657 CH1 delet
6	54	100.0	11	8	ADJ51618	Adj51618 CH1 delet
7	38	70.4	211	4	ABB68658	Abb68658 Drosophil
8	38	70.4	228	6	ABU50104	Abu50104 Protein e
9	36	66.7	147	7	ADM25854	Adm25854 Hyperther
10	36	66.7	161	5	ABP28993	Abp28993 Streptoco
11	36	66.7	165	5	ABP28646	Abp28646 Streptoco
12	36	66.7	165	8	ADK47927	Adk47927 Streptoco
13	36	66.7	165	8	ADV87871	Adv87871 Streptoco
14	36	66.7	165	8	ADV81323	Adv81323 Streptoco
15	36	66.7	165	8	ADV79124	Adv79124 Streptoco
16	36	66.7	166	5	ABP28647	Abp28647 Streptoco
17	36	66.7	188	2	AAW62717	Aaw62717 Streptoco
18	36	66.7	188	6	ABU02799	Abu02799 S. pneumo
19	36	66.7	189	8	ADR94919	Adr94919 Novel S.
20	36	66.7	189	9	AEA58789	Aea58789 Streptoco
21	35	64.8	163	5	ABP73897	Abp73897 Candida a
22	35	64.8	229	6	ABM67247	Abm67247 Photorhab
23	35	64.8	284	8	ABM81324	Abm81324 Tumour-as
24	35	64.8	332	4	AAU51057	Aau51057 Propionib
25	35	64.8	332	6	ABM47576	Abm47576 Propionib
26	35	64.8	396	6	ABP78925	Abp78925 N. gonorr
27	35	64.8	396	6	ABP76875	Abp76875 N. gonorr
28	35	64.8	396	8	ADS24847	Ads24847 Bacterial
29	35	64.8	569	4	AAB95893	Aab95893 Human pro
30	35	64.8	573	7	ADJ69587	Adj69587 Human hea
31	35	64.8	623	7	ABO77064	Abo77064 Pseudomon
32	35	64.8	638	5	AAO21853	Aao21853 Isoprenoi
33	34	63.0	10	2	AAY22361	Aay22361 TPO recep
34	34	63.0	10	3	AAB16984	Aab16984 TPO-mimet
35	34	63.0	10	5	ABB72870	Abb72870 TPO mimet
36	34	63.0	10	7	ADJ73021	Adj73021 TPO mimet
37	34	63.0	10	8	ADJ52656	Adj52656 CH1 delet
38	34	63.0	10	8	ADJ51617	Adj51617 CH1 delet
39	34	63.0	18	8	ABO53909	Abo53909 Human gen
40	34	63.0	71	5	ABP01520	Abp01520 Human ORF
41	34	63.0	221	6	ABU70526	Abu70526 Human adi
42	34	63.0	222	6	ABU70370	Abu70370 Human adi
43	34	63.0	222	6	ABU70476	Abu70476 Human adi
44	34	63.0	222	6	ABU70586	Abu70586 Human adi
45	34	63.0	295	5	ABG91460	Abg91460 Purine/py

ALIGNMENTS

RESULT 1
AAY22362
ID AAY22362 standard; peptide; 11 AA.
XX
AC AAY22362;
XX
DT 27-SEP-1999 (first entry)
XX
DE TPO receptor binding peptide sequence, SEQ ID NO. 13.
XX
KW TPO; thrombopoietin receptor; thrombopoietin agonist; thrombocytopaenia;
KW haematological disorder; therapy; bone marrow transfusion; diagnosis;
KW haematopoiesis; megakaryocyte expansion; thrombocyte regeneration.
XX
OS Synthetic.
XX
PN US5932546-A.
XX
PD 03-AUG-1999.
XX
PF 04-OCT-1996; 96US-00726464.
XX
PR 04-OCT-1996; 96US-00726464.
XX
PA (GLAX) GLAXO WELLCOME INC.
XX
PI Barrett RW, Wrighton NC, Duffin DJ, Wagstrom CR, Dower WJ;
PI Cwirla SE, Johnson SS;
XX
DR WPI; 1999-457122/38.
XX
PT New low molecular weight thrombopoietin agonists, particularly peptides,
PT for treatment of hematological disease and thrombocytopenia.
XX
PS Disclosure; Col 13-14; 36pp; English.
XX
CC This sequence represents a thrombopoietin (TPO) receptor (TR) binding
CC peptide of the invention. The peptide has: (i) a molecular weight below
CC 5000; and (ii) a binding affinity for TR, expressed as IC50, not over 10
CC mM. The peptides are used to treat conditions requiring a thrombopoietin
CC agonist, particularly haematological disorders or thrombocytopaenia,
CC especially resulting from chemotherapy, radiation therapy or bone marrow
CC transfusions. Also when labelled they may be used for diagnosis
CC (detecting TR on cells); for studying mechanisms of haematopoiesis; for
CC in vitro expansion of megakaryocytes and committed progenitor cells, and
CC for the development/identification of other TR agonists. The compounds
CC accelerate thrombocyte regeneration

XX
SQ Sequence 11 AA;

Query Match 100.0%; Score 54; DB 2; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.00044;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GVREVIVMHML 11
||||||||||||
Db 1 GVREVIVMHML 11

RESULT 2
AAB16985
ID AAB16985 standard; peptide; 11 AA.
XX
AC AAB16985;
XX
DT 31-OCT-2000 (first entry)
XX
DE TPO-mimetic peptide sequence SEQ ID NO:41.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US025044.
XX
PR 23-OCT-1998; 98US-0105371P.
PR 22-OCT-1999; 99US-00428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and pharmacologically
PT active peptides, useful for treating cancer and autoimmune diseases.
XX
PS Claim 19; Page 209; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and
CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC AAB18003 represent nucleotide and amino acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 11 AA;

Query Match 100.0%; Score 54; DB 3; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.00044;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GVREVIVMHML 11
||||||||||||
Db 1 GVREVIVMHML 11

RESULT 3
ABB72871
ID ABB72871 standard; peptide; 11 AA.
XX
AC ABB72871;
XX
DT 05-APR-2002 (first entry)
XX
DE TPO mimetic peptide SEQ ID NO:41.
XX
KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200183525-A2.
XX
PD 08-NOV-2001.
XX
PF 02-MAY-2001; 2001WO-US014310.
XX
PR 03-MAY-2000; 2000US-00563286.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
XX
DR WPI; 2002-130313/17.
XX
PT Novel vehicle-peptide molecule or its multimers useful for treating
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT diabetic retinopathy, obesity, sleep disorders and infertility.
XX
PS Claim 39; Page 43; 176pp; English.
XX
CC The present invention describes a vehicle-peptide molecule (I) or its
CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
CC neuroprotective activities. (I) can be used as a therapeutic or
CC prophylactic agent as well as for screening purposes. (I) is useful for
CC diagnosing diseases characterised by dysfunction of their associated
CC protein of interest, for identifying normal or abnormal proteins of
CC interest, as a part of diagnostic kit to detect the presence of their
CC proteins of interest in a biological sample. Additionally, (I) is useful
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
CC infertility, and neurological degenerative diseases. (I), comprising EPO-
CC mimetic compounds are useful for treating disorders characterised by low
CC red blood cell levels such as anaemia. The TPO-mimetic comprising
CC compounds are useful for treating conditions that involve an existing
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic

```
CC   tumour which result in thrombocytopaenia, systemic lupus erythematosus,
CC   and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC   represent amino acid and nucleic acid sequences used in the
CC   exemplification of the present invention
XX
SQ   Sequence 11 AA;

  Query Match             100.0%;  Score 54;  DB 5;  Length 11;
  Best Local Similarity   100.0%;  Pred. No. 0.00044;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GVREVIVMHML 11
              ||||||||||| |
Db          1 GVREVIVMHML 11


RESULT 4
ADJ73022
ID   ADJ73022 standard; peptide; 11 AA.
XX
AC   ADJ73022;
XX
DT   06-MAY-2004  (first entry)
XX
DE   TPO mimetic peptide sequence SeqID 476.
XX
KW   mimetic; CDR mimetibody; gene therapy; transgenic; immune;
KW   cardiovascular; infectious; malignant; neurologic disease; anaemia;
KW   immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;
KW   TPO.
XX
OS   Synthetic.
XX
PN   WO2003084477-A2.
XX
PD   16-OCT-2003.
XX
PF   24-MAR-2003; 2003WO-US009139.
XX
PR   29-MAR-2002; 2002US-0368791P.
XX
PA   (CENZ ) CENTOCOR INC.
XX
PI   Heavner GA,  Knight DM,  Scallon BJ,  Ghrayeb J;
XX
DR   WPI; 2003-804237/75.
XX
PT   New CDR mimetibody comprising a portion of a heavy or light chain
PT   variable region comprising human framework or ligand binding region,
PT   useful for preparing a composition for treating e.g., immune,
PT   cardiovascular or neurologic disease.
XX
PS   Disclosure; SEQ ID NO 476; 97pp; English.
XX
CC   This invention relates to novel mammalian CDR mimetibodies, specific
CC   portions or variants thereof. Specifically, it refers to an antibody
CC   fragment where a protein has been inserted into, or replaces a portion
CC   of, one or more CDR regions, such that each CDR mimetibody comprises at
CC   least one portion of a heavy chain or light chain variable region, which
CC   itself comprises at least one human framework region and at least one
CC   ligand binding region (LBR). The present invention describes human
CC   mimetibodies, including modified immunoglobulins and cleavage products
CC   that can be useful in gene therapy and the generation of transgenic
CC   plants and animals. Furthermore, the CDR mimetibody is useful for
CC   preparing compositions for modulating, treating or reducing the symptoms
CC   of immune, cardiovascular, infectious, malignant and/ or neurologic
CC   diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,
CC   cardiant, antimicrobial, cytostatic and neuroprotective activities. This
CC   peptide sequence is a TPO mimetic peptide sequence used to make a
CC   mimetibody of the invention.
XX
SQ   Sequence 11 AA;

  Query Match             100.0%;  Score 54;  DB 7;  Length 11;
  Best Local Similarity   100.0%;  Pred. No. 0.00044;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GVREVIVMHML 11
              ||||||||||| |
Db          1 GVREVIVMHML 11


RESULT 5
ADJ52657
ID   ADJ52657 standard; peptide; 11 AA.
XX
AC   ADJ52657;
XX
DT   06-MAY-2004  (first entry)
XX
DE   CH1 deleted mimetibody-related peptide SeqID476.
XX
KW   CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
KW   hypotensive; neuroprotective; nootropic; antibacterial; virucide;
KW   fungicide; gene therapy; immune disorder; cardiovascular disease;
KW   arrhythmia; hypertension; heart failure; neurodegenerative;
KW   multiple sclerosis; dementia; Alzheimer's disease; anaemia;
KW   cancerous condition; infectious disease; bacterial infection;
KW   viral infection; fungal infection.
XX
OS   Unidentified.
OS   Synthetic.
XX
PN   WO2004002417-A2.
XX
PD   08-JAN-2004.
XX
PF   27-JUN-2003; 2003WO-US020347.
XX
PR   28-JUN-2002; 2002US-0392431P.
XX
PA   (CENZ ) CENTOCOR INC.
XX
PI   Heavner GA,  Knight DM,  Ghrayeb J,  Scallon BJ,  Nesspor TC;
PI   Kutoloski KA;
XX
DR   WPI; 2004-082870/08.
XX
PT   New CH1-deleted mimetibody polypeptides and nucleic acids, useful for
PT   modulating, treating, alleviating, preventing an immune, cardiovascular,
PT   or neurodegenerative disease or disorder, anemia, cancer, or infectious
PT   diseases.
XX
PS   Claim 2; SEQ ID NO 476; 129pp; English.
XX
CC   This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC   which encode them), compositions, methods and uses. The invention may be
CC   useful for the development of compounds with an immunosuppressive,
CC   cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,
CC   antibacterial, virucide or fungicide activity. In addition, the disclosed
CC   sequences may prove useful for gene therapy. The CH1-deleted mimetibody
CC   is useful for diagnosing or treating a disease condition in a cell,
CC   tissue, organ or animal, specifically for modulating, treating,
CC   alleviating, preventing the incidence or reducing the symptoms of an
CC   immune, cardiovascular (for example arrhythmia, hypertension or heart
CC   failure), or neurodegenerative (for example multiple sclerosis, dementia
CC   or Alzheimer's disease) diseases or disorders, anaemia, cancerous
CC   conditions, or infectious diseases (for example bacterial, viral or
CC   fungal infection). The present sequence is that of a peptide which may be
CC   used during the creation of a mimetibody of the invention.
XX
SQ   Sequence 11 AA;

  Query Match             100.0%;  Score 54;  DB 8;  Length 11;
  Best Local Similarity   100.0%;  Pred. No. 0.00044;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 GVREVIVMHML 11
              ||||||||||
Db          1 GVREVIVMHML 11


RESULT 6
ADJ51618
ID   ADJ51618 standard; peptide; 11 AA.
XX
AC   ADJ51618;
XX
DT   06-MAY-2004  (first entry)
XX
DE   CH1 deleted mimetibody-related peptide SeqID476.
XX
KW   CH1 deleted mimetibody; osteopathic; cardiovascular-Gen;
KW   dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;
KW   gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;
KW   antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;
KW   ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;
KW   TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;
KW   dental disorder; oral disorder; dermatological disorder; ear disorder;
KW   nose disorder; throat disorder; endocrine disorder; metabolic disorder;
KW   gastrointestinal disorder; gynaecological disorder; hepatic disorder;
KW   obstetric disorder; haematologic disorder; immunological disorder;
KW   allergic disorder; infectious disorder; musculoskeletal disorder;
KW   oncological disorder; neurological disorder; nutritional disorder;
KW   ophthalmologic disorder; pediatric disorder; psychiatric disorder;
KW   renal disorder; pulmonary disorder.
XX
OS   Unidentified.
OS   Synthetic.
XX
PN   WO2004002424-A2.
XX
PD   08-JAN-2004.
XX
PF   30-JUN-2003; 2003WO-US020495.
XX
PR   28-JUN-2002; 2002US-0392431P.
PR   19-SEP-2002; 2002US-0412144P.
XX
PA   (CENZ ) CENTOCOR INC.
XX
PI   Heavner GA,  Knight DM,  Ghrayeb J,  Scallon BJ,  Nesspor TC;
PI   Kutoloski KA;
XX
DR   WPI; 2004-082872/08.
XX
PT   New CH1 deleted mimetibody polypeptide and nucleic acid, useful for
PT   diagnosing, preventing or treating cardiovascular, dermatologic,
PT   endocrine, gastrointestinal, gynecologic, infectious, neurologic and
PT   nutritional disorders.
XX
PS   Claim 14; SEQ ID NO 476; 123pp; English.
XX
CC   This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC   which encode them), compositions, methods and uses. The invention may be
CC   useful for the development of compounds with an osteopathic,
CC   cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,
CC   gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,
CC   immunomodulator, antiallergic, muscular-Gen, cytostatic,
CC   antiinflammatory, neuroleptic, ophthalmological, nephrotropic or
CC   respiratory-Gen activity acting as a tumour necrosis factor (TNF)-
CC   modulator or cytokine-agonist. The methods and compositions of the
CC   present invention are useful for the diagnosis, prevention and/or
CC   treatment of diseases or conditions associated with aberrant expression
CC   or activity of the CH1 deleted mimetibody, such as a bone or joint,
CC   cardiovascular, dental or oral, dermatological, ear, nose or throat,
CC   endocrine, metabolic, gastrointestinal, gynaecological, hepatic,
CC   obstetric, haematologic, immunological, allergic, infectious,
CC   musculoskeletal, oncological, neurological, nutritional, ophthalmologic,
CC   pediatric, psychiatric, renal or pulmonary disorders. The present
CC   sequence is that of a peptide which may be used during the creation of a
CC   mimetibody of the invention.
XX
SQ   Sequence 11 AA;

  Query Match             100.0%;  Score 54;  DB 8;  Length 11;
  Best Local Similarity   100.0%;  Pred. No. 0.00044;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GVREVIVMHML 11
              ||||||||||
Db          1 GVREVIVMHML 11


RESULT 7
ABB68658
ID   ABB68658 standard; protein; 211 AA.
XX
AC   ABB68658;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster polypeptide SEQ ID NO 32766.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US009231.
XX
PR   23-MAR-2000; 2000US-0191637P.
PR   11-JUL-2000; 2000US-00614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
DR   N-PSDB; ABL12761.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signaling and cell-cell
PT   interactions.
XX
PS   Disclosure; SEQ ID NO 32766; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC   ABB72072). The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 211 AA;

  Query Match             70.4%;  Score 38;  DB 4;  Length 211;
  Best Local Similarity   72.7%;  Pred. No. 22;
  Matches    8;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GVREVIVMHML 11
              ||  |||||:|
Db         90 GVATVIVMHVL 100
```

```
RESULT 8
ABU50104
ID   ABU50104 standard; protein; 228 AA.
XX
AC   ABU50104;
XX
DT   19-JUN-2003  (first entry)
XX
DE   Protein encoded by Prokaryotic essential gene #35631.
XX
KW   Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS   Yersinia pestis.
XX
PN   WO200277183-A2.
XX
PD   03-OCT-2002.
XX
PF   21-MAR-2002; 2002WO-US009107.
XX
PR   21-MAR-2001; 2001US-00815242.
PR   06-SEP-2001; 2001US-00948993.
PR   25-OCT-2001; 2001US-0342923P.
PR   08-FEB-2002; 2002US-00072851.
PR   06-MAR-2002; 2002US-0362699P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Wang L,  Zamudio C,  Malone C,  Haselbeck R,  Ohlsen KL,  Zyskind JW;
PI   Wall D,  Trawick JD,  Carr GJ,  Yamamoto R,  Forsyth RA,  Xu HH;
XX
DR   WPI; 2003-029926/02.
DR   N-PSDB; ACA53974.
XX
PT   New antisense nucleic acids, useful for identifying proteins or screening
PT   for homologous nucleic acids required for cellular proliferation to
PT   isolate candidate molecules for rational drug discovery programs.
XX
PS   Claim 25; SEQ ID NO 78028; 1766pp; English.
XX
CC   The invention relates to an isolated nucleic acid comprising any one of
CC   the 6213 antisense sequences given in the specification where expression
CC   of the nucleic acid inhibits proliferation of a cell. Also included are:
CC   (1) a vector comprising a promoter operably linked to the nucleic acid
CC   encoding a polypeptide whose expression is inhibited by the antisense
CC   nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC   polypeptide or its fragment whose expression is inhibited by the
CC   antisense nucleic acid; (4) an antibody capable of specifically binding
CC   the polypeptide;(5) producing the polypeptide; (6) inhibiting cellular
CC   proliferation or the activity of a gene in an operon required for
CC   proliferation; (7) identifying a compound that influences the activity of
CC   the gene product or that has an activity against a biological pathway
CC   required for proliferation, or that inhibits cellular proliferation; (8)
CC   identifying a gene required for cellular proliferation or the biological
CC   pathway in which a proliferation-required gene or its gene product lies
CC   or a gene on which the test compound that inhibits proliferation of an
CC   organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC   compound's activity; (11) a culture comprising strains in which the gene
CC   product is overexpressed or underexpressed; (12) determining the extent
CC   to which each of the strains is present in a culture or collection of
CC   strains; or (13) identifying the target of a compound that inhibits the
CC   proliferation of an organism. The antisense nucleic acids are useful for
CC   identifying proteins or screening for homologous nucleic acids required
CC   for cellular proliferation to isolate candidate molecules for rational
CC   drug discovery programs, or for screening homologous nucleic acids
CC   required for proliferation in cells other than S. aureus, S. typhimurium,
CC   K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC   the target prokaryotic essential genes. Note: The sequence data for this
CC   patent did not form part of the printed specification, but was obtained
CC   in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 228 AA;

  Query Match             70.4%;  Score 38;  DB 6;  Length 228;
  Best Local Similarity   77.8%;  Pred. No. 24;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GVREVIVMH 9
              | |||:|||
Db        147 GQREVVVMH 155


RESULT 9
ADM25854
ID   ADM25854 standard; protein; 147 AA.
XX
AC   ADM25854;
XX
DT   20-MAY-2004  (first entry)
XX
DE   Hyperthermophile Methanopyrus kandleri protein #460.
XX
KW   hyperthermophile; protein stability enhancement;
KW   protein activity enhancement.
XX
OS   Methanopyrus kandleri.
XX
PN   WO2003076575-A2.
XX
PD   18-SEP-2003.
XX
PF   04-MAR-2003; 2003WO-US006664.
XX
PR   04-MAR-2002; 2002US-0361742P.
PR   14-MAY-2002; 2002US-0380423P.
PR   16-SEP-2002; 2002US-0410974P.
XX
PA   (FIDE-) FIDELITY SYSTEMS INC.
PA   (MALY/) MALYKH A.
XX
PI   Slesarev AI,  Pavlov A,  Pavlova N,  Kozyavkin S;
XX
DR   WPI; 2003-748383/70.
DR   N-PSDB; ADM27081.
XX
PT   New isolated nucleic acids encoding any of about 1700 Methanopyrus
PT   kandleri proteins, and the encoded proteins, useful as a medicaments or
PT   as diagnostic agents.
XX
PS   Claim 31; SEQ ID NO 460; 1023pp; English.
XX
CC   The invention comprises the amino acid sequence of proteins from the
CC   hyperthermophile Methanopyrus kandleri, the invention also comprises the
CC   complete genome from Methanopyrus kandleri. The Methanopyrus kandleri
CC   proteins of the invention are useful for enhancing the stability and/or
CC   activity of other proteins. The Methanopyrus kandleri genome is useful in
CC   a variety of diagnostic and analytical methods. The present amino acid
CC   sequence represents a Methanopyrus kandleri protein of the invention.
XX
SQ   Sequence 147 AA;

  Query Match             66.7%;  Score 36;  DB 7;  Length 147;
  Best Local Similarity   70.0%;  Pred. No. 38;
  Matches    7;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 VREVIVMHML 11
              ||||:|:| |
Db          7 VREVVVIHEL 16


RESULT 10
ABP28993
ID   ABP28993 standard; protein; 161 AA.
XX
```

```
AC   ABP28993;
XX
DT   02-JUL-2002  (first entry)
XX
DE   Streptococcus polypeptide SEQ ID NO 7162.
XX
KW   Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KW   group A streptococcus; Streptococcus pyogenes; antibacterial;
KW   antiinflammatory; infection; vaccine; meningitis; gene therapy.
XX
OS   Streptococcus agalactiae.
XX
PN   WO200234771-A2.
XX
PD   02-MAY-2002.
XX
PF   29-OCT-2001; 2001WO-GB004789.
XX
PR   27-OCT-2000; 2000GB-00026333.
PR   24-NOV-2000; 2000GB-00028727.
PR   07-MAR-2001; 2001GB-00005640.
XX
PA   (CHIR-) CHIRON SPA.
PA   (GENO-) INST GENOMIC RES.
XX
PI   Telford J,  Masignani V,  Margarit Y RosI,  Grandi G,  Fraser C;
PI   Tettelin H;
XX
DR   WPI; 2002-352536/38.
DR   N-PSDB; ABN69624.
XX
PT   New Streptococcus protein for the treatment or prevention of infection or
PT   disease caused by Streptococcus bacteria, such as meningitis, and for
PT   detecting a compound that binds to the protein.
XX
PS   Claim 1; Page 3873; 4525pp; English.
XX
CC   The invention relates to a protein (ABP25413-ABP30895) from group B
CC   streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
CC   (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
CC   the specification. The proteins have antibacterial and antiinflammatory
CC   activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
CC   antibodies that bind (I) are used in the manufacture of medicaments for
CC   the treatment or prevention of infection or disease caused by
CC   Streptococcus bacteria, particularly S. agalactiae and S. pyrogenes.
CC   Nucleic acids encoding (I) are used to detect Streptococcus in a
CC   biological sample. (I) is used to determine whether a compound binds to
CC   (I). A composition comprising (I) or a nucleic acid encoding (I), may be
CC   used as a vaccine or diagnostic composition. The disease caused by
CC   Streptococcus that is prevented or treated may be meningitis. Nucleic
CC   acid encoding (I) may be used to recombinantly produce (I) and may be
CC   used in gene therapy. Antibodies to (I) are used for affinity
CC   chromatography, immunoassays, and distinguishing/identifying
CC   Streptococcus proteins
XX
SQ   Sequence 161 AA;

  Query Match             66.7%;  Score 36;  DB 5;  Length 161;
  Best Local Similarity   55.6%;  Pred. No. 42;
  Matches    5;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GVREVIVMH 9
              | ||::|:|
Db         81 GTREIVVLH 89


RESULT 11
ABP28646
ID   ABP28646 standard; protein; 165 AA.
XX
AC   ABP28646;
XX
DT   02-JUL-2002  (first entry)
XX
DE   Streptococcus polypeptide SEQ ID NO 6468.
XX
KW   Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KW   group A streptococcus; Streptococcus pyogenes; antibacterial;
KW   antiinflammatory; infection; vaccine; meningitis; gene therapy.
XX
OS   Streptococcus agalactiae.
XX
PN   WO200234771-A2.
XX
PD   02-MAY-2002.
XX
PF   29-OCT-2001; 2001WO-GB004789.
XX
PR   27-OCT-2000; 2000GB-00026333.
PR   24-NOV-2000; 2000GB-00028727.
PR   07-MAR-2001; 2001GB-00005640.
XX
PA   (CHIR-) CHIRON SPA.
PA   (GENO-) INST GENOMIC RES.
XX
PI   Telford J,  Masignani V,  Margarit Y RosI,  Grandi G,  Fraser C;
PI   Tettelin H;
XX
DR   WPI; 2002-352536/38.
DR   N-PSDB; ABN69277.
XX
PT   New Streptococcus protein for the treatment or prevention of infection or
PT   disease caused by Streptococcus bacteria, such as meningitis, and for
PT   detecting a compound that binds to the protein.
XX
PS   Claim 1; Page 3809; 4525pp; English.
XX
CC   The invention relates to a protein (ABP25413-ABP30895) from group B
CC   streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
CC   (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
CC   the specification. The proteins have antibacterial and antiinflammatory
CC   activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
CC   antibodies that bind (I) are used in the manufacture of medicaments for
CC   the treatment or prevention of infection or disease caused by
CC   Streptococcus bacteria, particularly S. agalactiae and S. pyrogenes.
CC   Nucleic acids encoding (I) are used to detect Streptococcus in a
CC   biological sample. (I) is used to determine whether a compound binds to
CC   (I). A composition comprising (I) or a nucleic acid encoding (I), may be
CC   used as a vaccine or diagnostic composition. The disease caused by
CC   Streptococcus that is prevented or treated may be meningitis. Nucleic
CC   acid encoding (I) may be used to recombinantly produce (I) and may be
CC   used in gene therapy. Antibodies to (I) are used for affinity
CC   chromatography, immunoassays, and distinguishing/identifying
CC   Streptococcus proteins
XX
SQ   Sequence 165 AA;

  Query Match             66.7%;  Score 36;  DB 5;  Length 165;
  Best Local Similarity   55.6%;  Pred. No. 43;
  Matches    5;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GVREVIVMH 9
              | ||::|:|
Db         81 GTREIVVLH 89


RESULT 12
ADK47927
ID   ADK47927 standard; protein; 165 AA.
XX
AC   ADK47927;
XX
DT   20-MAY-2004  (first entry)
XX
DE   Streptococcus pneumoniae protein, Seq ID No 4442.
XX
```

KW Antibacterial; Gene therapy; Vaccine; Streptococcus pneumoniae.
XX
OS Streptococcus pneumoniae.
XX
PN US6699703-B1.
XX
PD 02-MAR-2004.
XX
PF 26-MAY-2000; 2000US-00583110.
XX
PR 02-JUL-1997; 97US-0051553P.
PR 12-MAY-1998; 98US-0085131P.
PR 30-JUN-1998; 98US-00107433.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Doucette-Stamm L, Bush D, Zeng Q, Opperman T, Houseweart CE;
XX
DR WPI; 2004-212399/20.
DR N-PSDB; ADK45266.
XX
PT New nucleic acid molecules and polypeptides useful for diagnosing,
PT preventing and treating pathological conditions resulting from bacterial
PT infection, e.g. Streptococcus pneumoniae infection, and in drug
PT screening.
XX
PS Disclosure; SEQ ID NO 4442; 301pp; English.
XX
CC The invention relates to isolated Streptococcus pneumoniae nucleic acids
CC and polypeptides. The nucleic acids and proteins are useful for
CC diagnosing, preventing and treating pathological conditions resulting
CC from bacterial infection, such as S. pneumoniae infection. These may also
CC be used for drug screening procedures. The present sequence represents a
CC Streptococcus pneumoniae polypeptide of the invention. Note: The sequence
CC data for this patent did not appear in the printed specification but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 165 AA;

Query Match 66.7%; Score 36; DB 8; Length 165;
Best Local Similarity 55.6%; Pred. No. 43;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GVREVIVMH 9
| ||::|:|
Db 80 GTREIVVLH 88

RESULT 13
ADV87871
ID ADV87871 standard; protein; 165 AA.
XX
AC ADV87871;
XX
DT 24-FEB-2005 (first entry)
XX
DE Streptococcus agalactiae protein sequence, SEQ ID 265.
XX
KW Antibacterial; Vaccine; bacterial infection.
XX
OS Streptococcus agalactiae.
XX
PN FR2824074-A1.
XX
PD 31-OCT-2002.
XX
PF 26-APR-2001; 2001FR-00005642.
XX
PR 26-APR-2001; 2001FR-00005642.
XX
PA (INSP) INST PASTEUR.
PA (CNRS) CNRS CENT NAT RECH SCI.
XX
PI Glaser P, Rusniok C, Chevalier F, Frangeul L, Lalioui L;
PI Zouine M, Couve E, Buchrieser C, Poyart C, Trieu CP, Kunst F;
XX
DR WPI; 2004-101891/11.
XX
PT Genomic nucleotide sequences encoding polypeptides of Streptococcus
PT agalactiae for the development of vaccines, diagnostic tools, DNA chips
PT and identification of therapeutic targets.
XX
PS Claim 6; SEQ ID NO 265; 2687pp; French.
XX
CC The present invention relates to novel Streptococcus agalactiae
CC nucleotide sequences (I; ADV87607-ADV87745) and novel polypeptides (II;
CC ADV87746-ADV89950). The nucleotide sequences encode polypeptides of S.
CC agalactiae involved in the synthesis of amino acids, cell membranes,
CC intermediate (central) metabolism, energetic metabolism, fatty acid and
CC phospholipid metabolism, nucleotide metabolism including purines,
CC pyrimidines and/or nucleosides, regulatory functions, replication,
CC transcription, translation, protein transport, adaptation to atypical
CC conditions, sensitivity to medicines and/or analogues, functions related
CC to transposons, biosynthesis of cofactors, prosthetic groups and
CC transporters, cell membrane proteins and cellular machinery. (I) are
CC useful for the detection and/or amplification of nucleic acids.
CC Pharmaceutical composition comprising (I) or (II) are useful for
CC treatment of a bacterial S. agalactiae infection. Note: WO200292818A2 is
CC equivalent for the present basic patent FR2824074A1. WO200292818A2
CC contains 6617 sequence whereas the present patent only contains 2344
CC sequences.
XX
SQ Sequence 165 AA;

Query Match 66.7%; Score 36; DB 8; Length 165;
Best Local Similarity 55.6%; Pred. No. 43;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GVREVIVMH 9
| ||::|:|
Db 81 GTREIVVLH 89

RESULT 14
ADV81323
ID ADV81323 standard; protein; 165 AA.
XX
AC ADV81323;
XX
DT 24-FEB-2005 (first entry)
XX
DE Streptococcus agalactiae protein, SEQ ID 2464.
XX
KW Antibacterial; vaccine; bacterial infection.
XX
OS Streptococcus agalactiae.
XX
PN WO200292818-A2.
XX
PD 21-NOV-2002.
XX
PF 26-APR-2002; 2002WO-IB003059.
XX
PR 26-APR-2001; 2001FR-00005642.
XX
PA (INSP) INST PASTEUR.
PA (CNRS) CNRS CENT NAT RECH SCI.
XX
PI Glaser P, Rusniok C, Chevalier F, Frangeul L, Lalioui L;
PI Zouine M, Couve E, Buchrieser C, Poyart C, Trieu-Cuot P, Kunst F;
XX
DR WPI; 2004-101891/11.
XX
PT Genomic nucleotide sequences encoding polypeptides of Streptococcus
PT agalactiae for the development of vaccines, diagnostic tools, DNA chips

PT and identification of therapeutic targets.
XX
PS Claim 6; SEQ ID NO 2464; 439pp; French.
XX
CC The present invention relates to novel Streptococcus agalactiae
CC nucleotide sequences (I; ADV78860-ADV78998 and ADV83341-ADV85476) and
CC novel polypeptides (II; ADV78999-ADV81203 and ADV81205-ADV83340). The
CC nucleotide sequences encode polypeptides of S. agalactiae involved in the
CC synthesis of amino acids, cell membranes, intermediate (central)
CC metabolism, energetic metabolism, fatty acid and phospholipid metabolism,
CC nucleotide metabolism including purines, pyrimidines and/or nucleosides,
CC regulatory functions, replication, transcription, translation, protein
CC transport, adaptation to atypical conditions, sensitivity to medicines
CC and/or analogues, functions related to transposons, biosynthesis of
CC cofactors, prosthetic groups and transporters, cell membrane proteins and
CC cellular machinery. (I) are useful for the detection and/or amplification
CC of nucleic acids. Pharmaceutical composition comprising (I) or (II) are
CC useful for treatment of a bacterial S. agalactiae infection. The complete
CC genome of Streptococcus agalactiae is given in ADV81204. Note: The
CC present patent is an equivalent for the basic patent FR2824074A1, which
CC contains only 2344 sequences.
XX
SQ Sequence 165 AA;

Query Match 66.7%; Score 36; DB 8; Length 165;
Best Local Similarity 55.6%; Pred. No. 43;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GVREVIVMH 9
| ||::|:|
Db 81 GTREIVVLH 89

RESULT 15
ADV79124
ID ADV79124 standard; protein; 165 AA.
XX
AC ADV79124;
XX
DT 24-FEB-2005 (first entry)
XX
DE Streptococcus agalactiae protein, SEQ ID 265.
XX
KW Antibacterial; vaccine; bacterial infection.
XX
OS Streptococcus agalactiae.
XX
PN WO200292818-A2.
XX
PD 21-NOV-2002.
XX
PF 26-APR-2002; 2002WO-IB003059.
XX
PR 26-APR-2001; 2001FR-00005642.
XX
PA (INSP) INST PASTEUR.
PA (CNRS) CNRS CENT NAT RECH SCI.
XX
PI Glaser P, Rusniok C, Chevalier F, Frangeul L, Lalioui L;
PI Zouine M, Couve E, Buchrieser C, Poyart C, Trieu-Cuot P, Kunst F;
XX
DR WPI; 2004-101891/11.
XX
PT Genomic nucleotide sequences encoding polypeptides of Streptococcus
PT agalactiae for the development of vaccines, diagnostic tools, DNA chips
PT and identification of therapeutic targets.
XX
PS Claim 6; SEQ ID NO 265; 439pp; French.
XX
CC The present invention relates to novel Streptococcus agalactiae
CC nucleotide sequences (I; ADV78860-ADV78998 and ADV83341-ADV85476) and
CC novel polypeptides (II; ADV78999-ADV81203 and ADV81205-ADV83340). The
CC nucleotide sequences encode polypeptides of S. agalactiae involved in the
CC synthesis of amino acids, cell membranes, intermediate (central)
CC metabolism, energetic metabolism, fatty acid and phospholipid metabolism,
CC nucleotide metabolism including purines, pyrimidines and/or nucleosides,
CC regulatory functions, replication, transcription, translation, protein
CC transport, adaptation to atypical conditions, sensitivity to medicines
CC and/or analogues, functions related to transposons, biosynthesis of
CC cofactors, prosthetic groups and transporters, cell membrane proteins and
CC cellular machinery. (I) are useful for the detection and/or amplification
CC of nucleic acids. Pharmaceutical composition comprising (I) or (II) are
CC useful for treatment of a bacterial S. agalactiae infection. The complete
CC genome of Streptococcus agalactiae is given in ADV81204. Note: The
CC present patent is an equivalent for the basic patent FR2824074A1, which
CC contains only 2344 sequences.
XX
SQ Sequence 165 AA;

Query Match 66.7%; Score 36; DB 8; Length 165;
Best Local Similarity 55.6%; Pred. No. 43;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GVREVIVMH 9
| ||::|:|
Db 81 GTREIVVLH 89

Search completed: March 31, 2006, 16:22:25
Job time : 45.4129 secs

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: March 31, 2006, 16:22:51 ; Search time 6.8408 Seconds
(without alignments)
154.717 Million cell updates/sec

Title: US-10-609-217-41
Perfect score: 54
Sequence: 1 GVREVIVMHML 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	38	70.4	228	2	AE0325	urease accessory p
2	36	66.7	165	2	F95002	conserved hypothet
3	36	66.7	188	2	B97875	conserved hypothet
4	36	66.7	330	2	AC3370	6-aminohexanoate-d
5	35	64.8	110	2	D95267	hypothetical prote
6	35	64.8	274	2	F87648	dihydropteroate sy
7	35	64.8	384	2	E75095	hypothetical prote
8	35	64.8	507	1	S36805	cytochrome P450 71
9	35	64.8	638	2	E70528	probable dxs prote
10	34	63.0	131	2	T16905	hypothetical prote
11	34	63.0	163	2	H70771	hypothetical prote
12	34	63.0	295	2	C82828	glucose-1-phosphat
13	34	63.0	329	2	G83627	probable nucleosid
14	34	63.0	526	2	T01089	hypothetical prote
15	33	61.1	135	2	T46252	hypothetical prote
16	33	61.1	154	2	D70487	molybdopterin bios
17	33	61.1	221	2	T52622	probable peptidylp
18	33	61.1	232	2	H90214	hypothetical prote
19	33	61.1	259	2	C69122	hypothetical prote
20	33	61.1	288	2	G70606	hypothetical prote
21	33	61.1	476	2	JC4646	bone morphogenetic
22	33	61.1	533	2	A70464	D-3-phosphoglycera
23	33	61.1	719	2	A83127	probable toxin tra
24	33	61.1	2843	1	RBHUAP	adenomatous polypo
25	33	61.1	2845	2	I49505	adenomatous polypo
26	32	59.3	53	2	D90189	protein translatio
27	32	59.3	186	2	T22646	hypothetical prote
28	32	59.3	198	2	AB0112	probable 5-formylt
29	32	59.3	273	2	A70585	hypothetical prote
30	32	59.3	273	2	B70739	hypothetical prote
31	32	59.3	288	2	E64151	probable pyridoxal
32	32	59.3	293	2	D95954	probable glucose-1
33	32	59.3	293	2	D95859	probable dihydrodi
34	32	59.3	368	2	S01651	probable RNA-direc
35	32	59.3	396	2	S31959	acyl-[acyl-carrier
36	32	59.3	403	2	B83293	transaminase (EC 2
37	32	59.3	404	2	D95233	aminotransferase,
38	32	59.3	404	2	F98097	aspartate transami
39	32	59.3	416	2	AF1246	branched-chain alp
40	32	59.3	416	2	AB1609	branched-chain alp
41	32	59.3	469	2	D84857	hypothetical prote
42	32	59.3	510	1	I64162	mviN protein homol
43	32	59.3	970	2	A72028	preprotein translo
44	32	59.3	970	2	G86595	protein translocas
45	32	59.3	1096	2	A96607	protein disease re

ALIGNMENTS

RESULT 1
AE0325
urease accessory protein UreF [imported] - Yersinia pestis (strain CO92)
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C;Accession: AE0325
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AE0325
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-228 <KUR>
A;Cross-references: UNIPROT:Q9ZFR7; UNIPARC:UPI0000137DD8; GB:AL590842; PIDN:CAC92908.1;
C;Genetics:
A;Gene: ureF
C;Superfamily: urease accessory protein UreF

Query Match 70.4%; Score 38; DB 2; Length 228;
Best Local Similarity 77.8%; Pred. No. 3.8;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 GVREVIVMH 9
              | |||:|||
Db        147 GQREVVVMH 155
```

RESULT 2
F95002
conserved hypothetical protein SP0024 [imported] - Streptococcus pneumoniae (strain TIGR
C;Species: Streptococcus pneumoniae
C;Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
C;Accession: F95002
R;Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A;Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A;Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A;Reference number: A95000; MUID:21357209; PMID:11463916
A;Accession: F95002
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-165 <KUR>
A;Cross-references: UNIPROT:Q97TB7; UNIPARC:UPI00000C9C73; GB:AE005672; PIDN:AAK74215.1;
A;Experimental source: strain TIGR4
C;Genetics:
A;Gene: SP0024
C;Superfamily: Methanobacterium thermoautotrophicum carbonic anhydrase

Query Match 66.7%; Score 36; DB 2; Length 165;
Best Local Similarity 55.6%; Pred. No. 7;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 GVREVIVMH 9
              | ||::|:|
Db         80 GTREIVVLH 88
```

RESULT 3
B97875
conserved hypothetical protein spr0026 [imported] - Streptococcus pneumoniae (strain R6)
C;Species: Streptococcus pneumoniae
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C;Accession: B97875
R;Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszczak, L.; Burgett, S.; DeHoff, B.S.; E
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A;Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
A;Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A;Reference number: A97872; MUID:21429245; PMID:11544234
A;Accession: B97875
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-188 <KUR>
A;Cross-references: UNIPROT:Q8DRN9; UNIPARC:UPI00000E3384; GB:AE007317; PIDN:AAK98830.1;
C;Genetics:
A;Gene: spr0026

Query Match 66.7%; Score 36; DB 2; Length 188;
Best Local Similarity 55.6%; Pred. No. 8;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 GVREVIVMH 9
              | ||::|:|
Db        103 GTREIVVLH 111
```

RESULT 4
AC3370
6-aminohexanoate-dimer hydrolase (EC 3.5.1.46) [imported] - Brucella melitensis (strain
C;Species: Brucella melitensis
C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C;Accession: AC3370
R;DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
.; Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitens
A;Reference number: AD3252; PMID:11756688
A;Accession: AC3370
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-330 <KUR>
A;Cross-references: UNIPROT:Q8YH59; UNIPARC:UPI0000057E9B; GB:AE008917; PIDN:AAL52126.1;
A;Experimental source: strain 16M
C;Genetics:
A;Gene: BMEI0945
A;Map position: I
C;Keywords: hydrolase

Query Match 66.7%; Score 36; DB 2; Length 330;
Best Local Similarity 55.6%; Pred. No. 15;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 GVREVIVMH 9
              |:| ::|||
Db         37 GIRAIVVMH 45
```

RESULT 5
D95267
hypothetical protein SMa0087 [imported] - Sinorhizobium meliloti (strain 1021) magaplasmi
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C;Accession: D95267
R;Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bowse
.; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.;
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
A;Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti
A;Reference number: A95262; MUID:21396509; PMID:11481432
A;Accession: D95267
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-110 <KUR>
A;Cross-references: UNIPROT:Q931A0; UNIPARC:UPI00000CAF66; GB:AE006469; PIDN:AAK64702.1;
A;Experimental source: strain 1021, megaplasmid pSymA
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.C
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: SMa0087
A;Genome: plasmid

Query Match 64.8%; Score 35; DB 2; Length 110;
Best Local Similarity 77.8%; Pred. No. 7.3;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 GVREVIVMH 9
              ||||| |:|
Db         37 GVREVRVLH 45
```

RESULT 6
F87648
dihydropteroate synthase [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C;Accession: F87648
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.E
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolona
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: F87648
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-274 <STO>
A;Cross-references: UNIPROT:Q9A3I0; UNIPARC:UPI00000C79D1; GB:AE005673; NID:g13424906; PI
C;Genetics:
A;Gene: CC3224
C;Superfamily: dihydropteroate synthase; dihydropteroate synthase homology

Query Match 64.8%; Score 35; DB 2; Length 274;
Best Local Similarity 75.0%; Pred. No. 19;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```
Qy          4 EVIVMHML 11
              ||::||||
Db        128 EVVLMHML 135
```

RESULT 7
E75095
hypothetical protein PAB1606 - Pyrococcus abyssi (strain Orsay)
C;Species: Pyrococcus abyssi
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C;Accession: E75095

```
R;anonymous, Genoscope
submitted to the EMBL Data Library, July 1999
A;Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru
A;Reference number: A75001
A;Accession: E75095
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-384 <KAW>
A;Cross-references: UNIPROT:Q9UZJ1; UNIPARC:UPI00000633CA; GB:AJ248286; GB:AL096836; NID
A;Experimental source: strain Orsay
C;Genetics:
A;Gene: PAB1606

  Query Match             64.8%;  Score 35;  DB 2;  Length 384;
  Best Local Similarity   66.7%;  Pred. No. 28;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GVREVIVMH 9
              |: :|||||
Db        238 GISDVIVMH 246


RESULT 8
S36805
cytochrome P450 71A4 - eggplant
N;Contains: oxidoreductase (EC 1.-.-.-)
C;Species: Solanum melongena (eggplant, aubergine)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: S36805
R;Umemoto, N.; Kobayashi, O.; Ishizaki-Nishizawa, O.; Toguri, T.
FEBS Lett. 330, 169-173, 1993
A;Title: cDNAs sequences encoding cytochrome P450 (CYP71 family) from eggplant seedlings
A;Reference number: S36805; MUID:93374057; PMID:8365486
A;Accession: S36805
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-507 <UME>
A;Cross-references: UNIPROT:P37117; UNIPARC:UPI0000126CA3; EMBL:X70981; NID:g402223; PID
C;Genetics:
A;Gene: CYP71A4
C;Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
C;Keywords: chromoprotein; heme; iron; metalloprotein; oxidoreductase
F;305-470/Domain: cytochrome P450 homology <CYP>
F;448/Binding site: heme iron (Cys) (axial ligand) #status predicted

  Query Match             64.8%;  Score 35;  DB 1;  Length 507;
  Best Local Similarity   50.0%;  Pred. No. 37;
  Matches    5;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 VREVIVMHML 11
              || ::|:|:|
Db        131 VRSIVVLHLL 140


RESULT 9
E70528
probable dxs protein - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C;Accession: E70528
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: E70528
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-638 <COL>
A;Cross-references: UNIPROT:O07184; UNIPARC:UPI00001299E8; GB:Z96072; GB:AL123456; NID:g
A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: dxs
C;Superfamily: deoxyxylulose-5-phosphate synthase

  Query Match             64.8%;  Score 35;  DB 2;  Length 638;
  Best Local Similarity   63.6%;  Pred. No. 48;
  Matches    7;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 GVREVIVMHML 11
              ||||: | | |
Db        543 GVRELAVQHKL 553


RESULT 10
T16905
hypothetical protein T20B12.5 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C;Accession: T16905
R;Du, Z.
submitted to the EMBL Data Library, June 1994
A;Description: The sequence of C. elegans cosmid T20B12.
A;Reference number: S46772
A;Accession: T16905
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-131 <DUZ>
A;Cross-references: UNIPROT:P41845; UNIPARC:UPI000013BC50; EMBL:U10401; NID:g500713; PID
A;Experimental source: strain Bristol N2
C;Genetics:
A;Gene: CESP:T20B12.5

  Query Match             63.0%;  Score 34;  DB 2;  Length 131;
  Best Local Similarity   44.4%;  Pred. No. 14;
  Matches    4;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 VREVIVMHM 10
              :||::::||
Db         44 IREILLLHM 52


RESULT 11
H70771
hypothetical protein Rv1284 - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C;Accession: H70771
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: H70771
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-163 <COL>
A;Cross-references: UNIPROT:Q10612; UNIPARC:UPI000013A555; GB:Z73419; GB:AL123456; NID:g
A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: Rv1284
C;Superfamily: Methanobacterium thermoautotrophicum carbonic anhydrase

  Query Match             63.0%;  Score 34;  DB 2;  Length 163;
  Best Local Similarity   55.6%;  Pred. No. 18;
  Matches    5;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GVREVIVMH 9
              | ||:|::|
Db         79 GTREIILLH 87
```

RESULT 12
C82828
glucose-1-phosphate thymidylyltransferase XF0256 [imported] - Xylella fastidiosa (strain
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: C82828
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: C82828
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-295 <SIM>
A;Cross-references: UNIPROT:Q9PGP2; UNIPARC:UPI00000C234F; GB:AE003879; GB:AE003849; NID
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laign
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tsuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF0256
C;Superfamily: glucose-1-phosphate thymidylytransferase

Query Match 63.0%; Score 34; DB 2; Length 295;
Best Local Similarity 54.5%; Pred. No. 34;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GVREVIVMHML 11
|:|:|:|: ||
Db 48 GIRQVLVITML 58

RESULT 13
G83627
probable nucleoside hydrolase PA0143 [imported] - Pseudomonas aeruginosa (strain PAO1)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: G83627
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: G83627
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-329 <STO>
A;Cross-references: UNIPROT:Q9I6Y9; UNIPARC:UPI00000C4F5F; GB:AE004452; GB:AE004091; NID
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: PA0143
C;Superfamily: yaaF protein

Query Match 63.0%; Score 34; DB 2; Length 329;
Best Local Similarity 62.5%; Pred. No. 38;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GVREVIVM 8
|||||:::|
Db 161 GVREIVIM 168

RESULT 14
T01089
hypothetical protein T10P11.12 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004
C;Accession: T01089
R;Kaplan, N.; Johnson, D.; Schutz, K.; Gnoj, L.; Hoffman, J.; Till, S.; de la Bastide, M.
hi, M.; Martienssen, R.; Chen, E.Y.; Wilson, R.; McCombie, W.R.
submitted to the EMBL Data Library, November 1998
A;Description: Sequence of A. thaliana BAC T10P11 from chromosome IV.
A;Reference number: Z14248
A;Accession: T01089
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-526 <KAP>
A;Cross-references: UNIPROT:O49621; UNIPARC:UPI000000179A; EMBL:AC002330; NID:g2262135; P
A;Experimental source: cultivar Columbia
C;Genetics:
A;Map position: 4
A;Introns: 61/1; 145/3; 184/3; 205/1; 237/3; 262/3; 309/1; 322/3; 392/3; 415/2; 435/3
A;Note: T10P11.12
C;Superfamily: barley pathogen resistance protein Mlo

Query Match 63.0%; Score 34; DB 2; Length 526;
Best Local Similarity 60.0%; Pred. No. 62;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 2 VREVIVMHML 11
|:| ::||||
Db 84 VKENVLMHML 93

RESULT 15
T46252
hypothetical protein DKFZp761D051.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004
C;Accession: T46252
R;Ottenwaelder, B.; Obermaier, B.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000
A;Reference number: Z23031
A;Accession: T46252
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-135 <AAA>
A;Cross-references: UNIPROT:Q9NT69; UNIPARC:UPI0000070D1A; EMBL:AL137499
A;Experimental source: adult amygdala; clone DKFZp761D051
C;Genetics:
A;Note: DKFZp761D051.1

Query Match 61.1%; Score 33; DB 2; Length 135;
Best Local Similarity 54.5%; Pred. No. 23;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GVREVIVMHML 11
|||:::| | |
Db 118 GVRDMLVKHHL 128

Search completed: March 31, 2006, 16:37:15
Job time : 9.8408 secs

Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.

- If you encounter an accession number from an older search run against UniProt (results file extension **.rup)** that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

  http://www.pir.uniprot.org/database/archive.shtml

If you have any questions regarding this information or your results, please contact any STIC searcher.

When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: March 31, 2006, 16:09:36 ; Search time 41.209 Seconds
(without alignments)
188.328 Million cell updates/sec

Title: US-10-609-217-41
Perfect score: 54
Sequence: 1 GVREVIVMHML 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	40	74.1	404	2	Q86M90_ANOST	Q86m90 anopheles s
2	40	74.1	406	2	Q74C20_GEOSL	Q74c20 geobacter s
3	39	72.2	185	2	Q5YPK8_NOCFA	Q5ypk8 nocardia fa
4	39	72.2	214	2	Q4HX55_GIBZE	Q4hx55 gibberella
5	39	72.2	1753	2	Q8I5W3_PLAF7	Q8i5w3 plasmodium
6	38	70.4	211	2	Q9VCQ2_DROME	Q9vcq2 drosophila
7	38	70.4	217	2	Q4V5H3_DROME	Q4v5h3 drosophila
8	38	70.4	228	1	UREF_YEREN	P42870 yersinia en
9	38	70.4	228	1	UREF_YERPE	Q9zfr7 yersinia pe
10	38	70.4	228	1	UREF_YERPS	P52318 yersinia ps
11	38	70.4	228	2	Q6UR31_YERRO	Q6ur31 yersinia ro
12	38	70.4	228	2	Q6UR40_YERMO	Q6ur40 yersinia mo
13	38	70.4	228	2	Q6UR49_YERKR	Q6ur49 yersinia kr
14	38	70.4	228	2	Q6UR66_YERFR	Q6ur66 yersinia fr
15	38	70.4	228	2	Q6UR74_YERBE	Q6ur74 yersinia be
16	38	70.4	228	2	Q6UR82_YERAL	Q6ur82 yersinia al
17	38	70.4	529	2	Q7PZ68_ANOGA	Q7pz68 anopheles g
18	38	70.4	2513	2	Q5PAX1_ANAMM	Q5pax1 anaplasma m
19	37	68.5	82	2	Q971Q4_SULTO	Q971q4 sulfolobus
20	37	68.5	227	2	Q4J8P4_SULAC	Q4j8p4 sulfolobus
21	37	68.5	270	1	RECX_BACAN	Q81yw4 bacillus an
22	37	68.5	270	1	RECX_BACHK	Q6hnu0 bacillus th
23	37	68.5	398	1	METK_BRAJA	Q89hp5 bradyrhizob
24	37	68.5	664	2	Q4N5M4_THEPA	Q4n5m4 theileria p
25	37	68.5	1076	2	Q4UDW9_THEAN	Q4udw9 theileria a
26	36	66.7	147	2	Q8TY46_METKA	Q8ty46 methanopyru
27	36	66.7	165	2	Q8E286_STRA5	Q8e286 streptococc
28	36	66.7	165	2	Q8E7P4_STRA3	Q8e7p4 streptococc
29	36	66.7	165	2	Q97TB7_STRPN	Q97tb7 streptococc
30	36	66.7	165	2	Q9A1K0_STRPY	Q9a1k0 streptococc
31	36	66.7	165	2	Q5LXZ3_STRT1	Q5lxz3 streptococc
32	36	66.7	165	2	Q5M2J7_STRT2	Q5m2j7 streptococc
33	36	66.7	165	2	Q8P2Q4_STRP8	Q8p2q4 streptococc
34	36	66.7	165	2	Q8K8Q0_STRP3	Q8k8q0 streptococc
35	36	66.7	166	2	Q5XDZ6_STRP6	Q5xdz6 streptococc
36	36	66.7	188	2	Q8DRN9_STRR6	Q8drn9 streptococc
37	36	66.7	330	2	Q8YH59_BRUME	Q8yh59 brucella me
38	36	66.7	440	2	Q57D89_BRUAB	Q57d89 brucella ab
39	36	66.7	440	2	Q8G0P7_BRUSU	Q8g0p7 brucella su
40	36	66.7	499	2	Q5QY17_IDILO	Q5qy17 idiomarina
41	36	66.7	834	1	ATG1_CANAL	Q5a649 candida alb
42	35	64.8	109	2	Q5QC10_9CAUD	Q5qc10 enterobacte
43	35	64.8	110	2	Q931A0_RHIME	Q931a0 rhizobium m
44	35	64.8	163	2	Q5A207_CANAL	Q5a207 candida alb
45	35	64.8	166	2	Q4III8_GIBZE	Q4iii8 gibberella

ALIGNMENTS

RESULT 1
Q86M90_ANOST
ID Q86M90_ANOST PRELIMINARY; PRT; 404 AA.
AC Q86M90;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE SG1D salivary protein precursor.
OS Anopheles stephensi (Indo-Pakistan malaria mosquito).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Culicidae;
OC Anophelinae; Anopheles.
OX NCBI_TaxID=30069;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22710796; PubMed=12826099; DOI=10.1016/S0965-1748(03)00067-5;
RA Valenzuela J.G., Francischetti I.M.B., Pham V.M., Garfield M.K.,
RA Ribeiro J.M.C.;
RT "Exploring the salivary gland transcriptome and proteome of the
RT Anopheles stephensi mosquito.";
RL Insect Biochem. Mol. Biol. 33:717-732(2003).
DR EMBL; AY226459; AAO74845.1; -; mRNA.
KW Signal.
FT SIGNAL 1 21 Potential.
SQ SEQUENCE 404 AA; 46840 MW; 4DC31C9F65704105 CRC64;

Query Match 74.1%; Score 40; DB 2; Length 404;
Best Local Similarity 63.6%; Pred. No. 24;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 GVREVIVMHML 11
              || ||:|:|:|
Db        175 GVNEVLVLHVL 185
```

RESULT 2
Q74C20_GEOSL
ID Q74C20_GEOSL PRELIMINARY; PRT; 406 AA.
AC Q74C20;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Capsule polysaccharide export protein, putative.
GN OrderedLocusNames=GSU1855;
OS Geobacter sulfurreducens.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfuromonadales;
OC Geobacteraceae; Geobacter.
OX NCBI_TaxID=35554;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PCA / ATCC 51573;
RX PubMed=14671304; DOI=10.1126/science.1088727;
RA Methe B.A., Nelson K.E., Eisen J.A., Paulsen I.T., Nelson W.C.,
RA Heidelberg J.F., Wu D., Wu M., Ward N.L., Beanan M.J., Dodson R.J.,

```
RA   Madupu R., Brinkac L.M., Daugherty S.C., DeBoy R.T., Durkin A.S.,
RA   Gwinn M.L., Kolonay J.F., Sullivan S.A., Haft D.H., Selengut J.,
RA   Davidsen T.M., Zafar N., White O., Tran B., Romero C., Forberger H.A.,
RA   Weidman J.F., Khouri H.M., Feldblyum T.V., Utterback T.R.,
RA   Van Aken S.E., Lovley D.R., Fraser C.M.;
RT   "Genome of Geobacter sulfurreducens: metal reduction in subsurface
RT   environments.";
RL   Science 302:1967-1969(2003).
DR   EMBL; AE017180; AAR35232.1; -; Genomic_DNA.
DR   TIGR; GSU1855; -.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0009103; P:lipopolysaccharide biosynthesis; IEA.
DR   InterPro; IPR003856; LPS_Wzz_MPA.
DR   Pfam; PF02706; Wzz; 1.
KW   Complete proteome.
SQ   SEQUENCE   406 AA;  45285 MW;  861A484CE8258AFF CRC64;

  Query Match             74.1%;  Score 40;  DB 2;  Length 406;
  Best Local Similarity   63.6%;  Pred. No. 24;
  Matches    7;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GVREVIVMHML 11
              |:  ||:||||
Db        374 GILSVIIMHML 384


RESULT 3
Q5YPK8_NOCFA
ID   Q5YPK8_NOCFA PRELIMINARY;      PRT;   185 AA.
AC   Q5YPK8;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Hypothetical protein.
GN   OrderedLocusNames=nfa50310;
OS   Nocardia farcinica.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Nocardiaceae; Nocardia.
OX   NCBI_TaxID=37329;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=IFM 10152;
RX   PubMed=15466710; DOI=10.1073/pnas.0406410101;
RA   Ishikawa J., Yamashita A., Mikami Y., Hoshino Y., Kurita H., Hotta K.,
RA   Shiba T., Hattori M.;
RT   "The complete genomic sequence of Nocardia farcinica IFM 10152.";
RL   Proc. Natl. Acad. Sci. U.S.A. 101:14925-14930(2004).
DR   EMBL; AP006618; BAD59883.1; -; Genomic_DNA.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   185 AA;  19914 MW;  8E6AFE6C846A9171 CRC64;

  Query Match             72.2%;  Score 39;  DB 2;  Length 185;
  Best Local Similarity   77.8%;  Pred. No. 18;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GVREVIVMH 9
              |||||:| |
Db         60 GVREVVVFH 68


RESULT 4
Q4HX55_GIBZE
ID   Q4HX55_GIBZE PRELIMINARY;      PRT;   214 AA.
AC   Q4HX55;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Hypothetical protein.
GN   ORFNames=FG10453.1;
OS   Gibberella zeae PH-1.
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC   Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
OX   NCBI_TaxID=229533;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=PH-1;
RA   Birren B., Nusbaum C., Abouelleil A., Allen N., Anderson S.,
RA   Arachchi H.M., Barna N., Bastien V., Bloom T., Boguslavkiy L.,
RA   Boukhgalter B., Butler J., Calvo S.E., Camarata J., Chang J.,
RA   Choepel Y., Collymore A., Cook A., Cooke P., Corum B., DeArellano K.,
RA   Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
RA   Erickson J., Faro S., Ferreira P., FitzGerald M., Gage D., Galagan J.,
RA   Gardyna S., Gnerre S., Graham L., Grand-Pierre N., Hafez N.,
RA   Hagopian D., Hagos B., Hall J., Horton L., Hulme W., Iliev I.,
RA   Jaffe D., Johnson R., Jones C., Kamal M., Kamat A., Karatas A.,
RA   Kells C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
RA   Ma L.-J., Mabbitt R., MacLean C., Macdonald P., Major J., Manning J.,
RA   Matthews C., Mauceli E., McCarthy M., Meldrim J., Meneus L.,
RA   Mihova T., Mlenga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
RA   Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neil D.,
RA   Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,
RA   Rachupka A., Ramasamy U., Raymond C., Retta R., Rise C., Rogov P.,
RA   Roman J., Schauer S., Schupback R., Seaman S., Severy P., Smirnov S.,
RA   Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,
RA   Talamas J., Tesfaye S., Theodore J., Topham K., Travers M.,
RA   Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,
RA   Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
RA   Lander E.;
RT   "Fusarium graminearum genome sequence.";
RL   Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; AACM01000435; EAA68227.1; -; Genomic_DNA.
KW   Hypothetical protein.
SQ   SEQUENCE   214 AA;  23847 MW;  E710E02437B5DF54 CRC64;

  Query Match             72.2%;  Score 39;  DB 2;  Length 214;
  Best Local Similarity   66.7%;  Pred. No. 20;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GVREVIVMH 9
              | ||:|:||
Db        121 GTREIIIMH 129


RESULT 5
Q8I5W3_PLAF7
ID   Q8I5W3_PLAF7 PRELIMINARY;      PRT;  1753 AA.
AC   Q8I5W3;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical protein.
GN   ORFNames=PFL0425c;
OS   Plasmodium falciparum (isolate 3D7).
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=36329;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22255705; PubMed=12368864; DOI=10.1038/nature01097;
RA   Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA   Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA   Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA   Chan M.-S., Nene V., Shallom S.J., Suh B., Peterson J., Angiuoli S.,
RA   Pertea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA   Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA   McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA   Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA   Fraser C.M., Barrell B.G.;
RT   "Genome sequence of the human malaria parasite Plasmodium
RT   falciparum.";
RL   Nature 419:498-511(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RA   Hyman R.W., Fung E., Conway A., Kurdi O., Mao J., Miranda M.,
```

```
RA   Nakao B., Rowley D., Tamaki T., Wang F., Davis R.W.;
RL   Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE014845; AAN36174.1; -; Genomic_DNA.
KW   Hypothetical protein.
SQ   SEQUENCE   1753 AA;  208672 MW;  FF0357DC8958A38A CRC64;

  Query Match             72.2%;  Score 39;  DB 2;  Length 1753;
  Best Local Similarity   60.0%;  Pred. No. 1.6e+02;
  Matches    6;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 VREVIVMHML 11
              ::|||::|||
Db        681 IKEVIIIHML 690


RESULT 6
Q9VCQ2_DROME
ID   Q9VCQ2_DROME PRELIMINARY;      PRT;   211 AA.
AC   Q9VCQ2;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   CG13829-PA.
GN   Name=CG13829; ORFNames=CG13829;
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22426065; PubMed=12537568;
RA   Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA   Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA   George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,
RA   Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA   Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA   Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT   "Finishing a whole-genome shotgun: release 3 of the Drosophila
RT   melanogaster euchromatic genome sequence.";
RL   Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22426070; PubMed=12537573;
RA   Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J.W., Svirskas R.,
RA   Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA   Ashburner M., Celniker S.E.;
RT   "The transposable elements of the Drosophila melanogaster euchromatin:
RT   a genomics perspective.";
RL   Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).
RN   [4]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22426069; PubMed=12537572;
RA   Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA   Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA   Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA   Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA   Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA   Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA   Lewis S.E.;
RT   "Annotation of the Drosophila melanogaster euchromatic genome: a
RT   systematic review.";
RL   Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN   [5]
RP   NUCLEOTIDE SEQUENCE.
RG   Berkeley Drosophila Genome Project;
RA   Celniker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,
RA   Hoskins R., Stapleton M., Pacleb J., Park S., Svirskas R., Smith E.,
RA   Yu C., Rubin G.;
RT   "Drosophila melanogaster release 4 sequence.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   NUCLEOTIDE SEQUENCE.
RG   FlyBase;
RL   Submitted (MAR-2005) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE003742; AAF56105.1; -; Genomic_DNA.
DR   Ensembl; CG13829; Drosophila melanogaster.
DR   FlyBase; FBgn0039059; CG13829.
SQ   SEQUENCE   211 AA;  23509 MW;  C9808024BE54DD5F CRC64;

  Query Match             70.4%;  Score 38;  DB 2;  Length 211;
  Best Local Similarity   72.7%;  Pred. No. 32;
  Matches    8;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GVREVIVMHML 11
              ||  |||||:|
Db         90 GVATVIVMHVL 100


RESULT 7
Q4V5H3_DROME
ID   Q4V5H3_DROME PRELIMINARY;      PRT;   217 AA.
AC   Q4V5H3;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   IP07660p (Fragment).
GN   Name=CG13829;
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
```

```
RA   Stapleton M., Carlson J., Chavez C., Frise E., George R., Pacleb J.,
RA   Park S., Wan K., Yu C., Celniker S.;
RL   Submitted (MAY-2005) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BT022683; AAY55099.1; -; mRNA.
FT   NON_TER       1      1
SQ   SEQUENCE   217 AA;  24271 MW;  F523AE4F2D3ECE22 CRC64;

  Query Match             70.4%;  Score 38;  DB 2;  Length 217;
  Best Local Similarity   72.7%;  Pred. No. 33;
  Matches    8;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GVREVIVMHML 11
              ||  |||||:|
Db         96 GVATVIVMHVL 106


RESULT 8
UREF_YEREN
ID   UREF_YEREN     STANDARD;      PRT;   228 AA.
AC   P42870;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Urease accessory protein ureF.
GN   Name=ureF;
OS   Yersinia enterocolitica.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Yersinia.
OX   NCBI_TaxID=630;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RC   STRAIN=A2635 / Serotype O:8;
RX   MEDLINE=94320783; PubMed=8045421; DOI=10.1016/0378-1119(94)90318-2;
RA   de Koning-Ward T.F., Ward A.C., Robins-Browne R.M.;
RT   "Characterisation of the urease-encoding gene complex of Yersinia
RT   enterocolitica.";
RL   Gene 145:25-32(1994).
CC   -!- FUNCTION: Probably facilitates nickel incorporation.
CC   -!- SIMILARITY: Belongs to the ureF family.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; L24101; AAA50998.1; -; Genomic_DNA.
DR   InterPro; IPR002639; UreF.
DR   Pfam; PF01730; UreF; 1.
DR   PIRSF; PIRSF009467; Ureas_acces_UreF; 1.
KW   Nickel.
SQ   SEQUENCE   228 AA;  25041 MW;  0EB2536245EBFB34 CRC64;

  Query Match             70.4%;  Score 38;  DB 1;  Length 228;
  Best Local Similarity   77.8%;  Pred. No. 34;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GVREVIVMH 9
              | |||:|||
Db        147 GQREVVVMH 155


RESULT 9
UREF_YERPE
ID   UREF_YERPE     STANDARD;      PRT;   228 AA.
AC   Q9ZFR7;
DT   30-MAY-2000 (Rel. 39, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-MAY-2005 (Rel. 47, Last annotation update)
DE   Urease accessory protein ureF.
GN   Name=ureF; OrderedLocusNames=YPO2669, y1241, YP2470;
OS   Yersinia pestis.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Yersinia.
OX   NCBI_TaxID=632;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=6/69M;
RA   Sebbane F., Devalckenaere A., Simonet M.;
RT   "Characterization of the urease locus from Yersinia pestis.";
RL   Submitted (SEP-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC   STRAIN=CO-92 / Biovar Orientalis;
RX   MEDLINE=21470413; PubMed=11586360; DOI=10.1038/35097083;
RA   Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA   Prentice M.B., Sebaihia M., James K.D., Churcher C.M., Mungall K.L.,
RA   Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.-M.,
RA   Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA   Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA   Leather S., Moule S., Oyston P.C.F., Quail M.A., Rutherford K.M.,
RA   Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
RT   "Genome sequence of Yersinia pestis, the causative agent of plague.";
RL   Nature 413:523-527(2001).
RN   [3]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC   STRAIN=KIM5 / Biovar Mediaevalis;
RX   MEDLINE=22137863; PubMed=12142430;
RX   DOI=10.1128/JB.184.16.4601-4611.2002;
RA   Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA   Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA   Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA   Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
RA   Perry R.D.;
RT   "Genome sequence of Yersinia pestis KIM.";
RL   J. Bacteriol. 184:4601-4611(2002).
RN   [4]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC   STRAIN=91001 / Biovar Mediaevalis;
RX   PubMed=15368893;
RA   Song Y., Tong Z., Wang J., Wang L., Guo Z., Han Y., Zhang J., Pei D.,
RA   Zhou D., Qin H., Pang X., Han Y., Zhai J., Li M., Cui B., Qi Z.,
RA   Jin L., Dai R., Chen F., Li S., Ye C., Du Z., Lin W., Wang J., Yu J.,
RA   Yang H., Wang J., Huang P., Yang R.;
RT   "Complete genome sequence of Yersinia pestis strain 91001, an isolate
RT   avirulent to humans.";
RL   DNA Res. 11:179-197(2004).
CC   -!- FUNCTION: Probably facilitates nickel incorporation.
CC   -!- SIMILARITY: Belongs to the ureF family.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; AF095636; AAC78636.1; -; Genomic_DNA.
DR   EMBL; AJ414153; CAC92908.1; -; Genomic_DNA.
DR   EMBL; AE013727; AAM84816.1; -; Genomic_DNA.
DR   EMBL; AE017136; AAS62670.1; -; Genomic_DNA.
DR   PIR; AE0325; AE0325.
DR   InterPro; IPR002639; UreF.
DR   Pfam; PF01730; UreF; 1.
DR   PIRSF; PIRSF009467; Ureas_acces_UreF; 1.
KW   Complete proteome; Nickel.
FT   CONFLICT     70     70       V -> M (in Ref. 1).
SQ   SEQUENCE   228 AA;  25037 MW;  AB1631F780AD54AC CRC64;

  Query Match             70.4%;  Score 38;  DB 1;  Length 228;
  Best Local Similarity   77.8%;  Pred. No. 34;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GVREVIVMH 9
              | |||:|||
Db        147 GQREVVVMH 155
```

```
RESULT 10
UREF_YERPS
ID   UREF_YERPS     STANDARD;      PRT;   228 AA.
AC   P52318; Q667Q2;
DT   01-OCT-1996 (Rel. 34, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   10-MAY-2005 (Rel. 47, Last annotation update)
DE   Urease accessory protein ureF.
GN   Name=ureF; OrderedLocusNames=YPTB2940;
OS   Yersinia pseudotuberculosis.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Yersinia.
OX   NCBI_TaxID=633;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=IP 2777;
RX   MEDLINE=97270511; PubMed=9125594;
RA   Riot B., Berche P., Simonet M.;
RT   "Urease is not involved in the virulence of Yersinia
RT   pseudotuberculosis in mice.";
RL   Infect. Immun. 65:1985-1990(1997).
RN   [2]
RP   SEQUENCE REVISION TO 5-9 AND 168-171.
RA   Riot B.;
RL   Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC   STRAIN=IP32953 / Serotype I;
RX   PubMed=15358858; DOI=10.1073/pnas.0404012101;
RA   Chain P.S.G., Carniel E., Larimer F.W., Lamerdin J., Stoutland P.O.,
RA   Regala W.M., Georgescu A.M., Vergez L.M., Land M.L., Motin V.L.,
RA   Brubaker R.R., Fowler J., Hinnebusch J., Marceau M., Medigue C.,
RA   Simonet M., Chenal-Francisque V., Souza B., Dacheux D., Elliott J.M.,
RA   Derbise A., Hauser L.J., Garcia E.;
RT   "Insights into the evolution of Yersinia pestis through whole-genome
RT   comparison with Yersinia pseudotuberculosis.";
RL   Proc. Natl. Acad. Sci. U.S.A. 101:13826-13831(2004).
CC   -!- FUNCTION: Probably facilitates nickel incorporation.
CC   -!- SIMILARITY: Belongs to the ureF family.
CC   ----------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ----------------------------------------------------------------------
DR   EMBL; U40842; AAA87856.2; -; Genomic_DNA.
DR   EMBL; BX936398; CAH22178.1; -; Genomic_DNA.
DR   InterPro; IPR002639; UreF.
DR   Pfam; PF01730; UreF; 1.
DR   PIRSF; PIRSF009467; Ureas_acces_UreF; 1.
KW   Complete proteome; Nickel.
SQ   SEQUENCE   228 AA;  25009 MW;  212D225507D42E67 CRC64;

  Query Match             70.4%;  Score 38;  DB 1;  Length 228;
  Best Local Similarity   77.8%;  Pred. No. 34;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GVREVIVMH 9
              | |||:|||
Db        147 GQREVVVMH 155


RESULT 11
Q6UR31_YERRO
ID   Q6UR31_YERRO PRELIMINARY;      PRT;   228 AA.
AC   Q6UR31;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   UreF.
GN   Name=ureF;
OS   Yersinia rohdei.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Yersinia.
OX   NCBI_TaxID=29485;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Sebbane F., Lemaitre N., Simonet M.;
RL   Submitted (AUG-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY363686; AAR15139.1; -; Genomic_DNA.
DR   GO; GO:0016151; F:nickel ion binding; IEA.
DR   GO; GO:0006807; P:nitrogen compound metabolism; IEA.
DR   InterPro; IPR002639; UreF.
DR   Pfam; PF01730; UreF; 1.
DR   PIRSF; PIRSF009467; Ureas_acces_UreF; 1.
SQ   SEQUENCE   228 AA;  25031 MW;  FBFD54D765ECBBC7 CRC64;

  Query Match             70.4%;  Score 38;  DB 2;  Length 228;
  Best Local Similarity   77.8%;  Pred. No. 34;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GVREVIVMH 9
              | |||:|||
Db        147 GQREVVVMH 155


RESULT 12
Q6UR40_YERMO
ID   Q6UR40_YERMO PRELIMINARY;      PRT;   228 AA.
AC   Q6UR40;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   UreF.
GN   Name=ureF;
OS   Yersinia mollaretii.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Yersinia.
OX   NCBI_TaxID=33060;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Sebbane F., Lemaitre N., Simonet M.;
RL   Submitted (AUG-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY363685; AAR15130.1; -; Genomic_DNA.
DR   GO; GO:0016151; F:nickel ion binding; IEA.
DR   GO; GO:0006807; P:nitrogen compound metabolism; IEA.
DR   InterPro; IPR002639; UreF.
DR   Pfam; PF01730; UreF; 1.
DR   PIRSF; PIRSF009467; Ureas_acces_UreF; 1.
SQ   SEQUENCE   228 AA;  25054 MW;  ECDA3BF0B9816592 CRC64;

  Query Match             70.4%;  Score 38;  DB 2;  Length 228;
  Best Local Similarity   77.8%;  Pred. No. 34;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GVREVIVMH 9
              | |||:|||
Db        147 GQREVVVMH 155


RESULT 13
Q6UR49_YERKR
ID   Q6UR49_YERKR PRELIMINARY;      PRT;   228 AA.
AC   Q6UR49;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   UreF.
GN   Name=ureF;
OS   Yersinia kristensenii.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Yersinia.
OX   NCBI_TaxID=28152;
RN   [1]
```

```
RP   NUCLEOTIDE SEQUENCE.
RA   Sebbane F., Lemaitre N., Simonet M.;
RL   Submitted (AUG-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY363684; AAR15121.1; -; Genomic_DNA.
DR   GO; GO:0016151; F:nickel ion binding; IEA.
DR   GO; GO:0006807; P:nitrogen compound metabolism; IEA.
DR   InterPro; IPR002639; UreF.
DR   Pfam; PF01730; UreF; 1.
DR   PIRSF; PIRSF009467; Ureas_acces_UreF; 1.
SQ   SEQUENCE   228 AA;  25106 MW;  BEFCBC531AA5763F CRC64;

  Query Match             70.4%;  Score 38;  DB 2;  Length 228;
  Best Local Similarity   77.8%;  Pred. No. 34;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GVREVIVMH 9
              | |||:|||
Db        147 GQREVVVMH 155


RESULT 14
Q6UR66_YERFR
ID   Q6UR66_YERFR PRELIMINARY;      PRT;   228 AA.
AC   Q6UR66;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   UreF.
GN   Name=ureF;
OS   Yersinia frederiksenii.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Yersinia.
OX   NCBI_TaxID=29484;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Sebbane F., Lemaitre N., Simonet M.;
RL   Submitted (AUG-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY363682; AAR15104.1; -; Genomic_DNA.
DR   GO; GO:0016151; F:nickel ion binding; IEA.
DR   GO; GO:0006807; P:nitrogen compound metabolism; IEA.
DR   InterPro; IPR002639; UreF.
DR   Pfam; PF01730; UreF; 1.
DR   PIRSF; PIRSF009467; Ureas_acces_UreF; 1.
SQ   SEQUENCE   228 AA;  25009 MW;  4C171AEC9D669422 CRC64;

  Query Match             70.4%;  Score 38;  DB 2;  Length 228;
  Best Local Similarity   77.8%;  Pred. No. 34;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GVREVIVMH 9
              | |||:|||
Db        147 GQREVVVMH 155


RESULT 15
Q6UR74_YERBE
ID   Q6UR74_YERBE PRELIMINARY;      PRT;   228 AA.
AC   Q6UR74;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   UreF.
GN   Name=ureF;
OS   Yersinia bercovieri.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Yersinia.
OX   NCBI_TaxID=634;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Sebbane F., Lemaitre N., Simonet M.;
RL   Submitted (AUG-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY363681; AAR15096.1; -; Genomic_DNA.
DR   GO; GO:0016151; F:nickel ion binding; IEA.
DR   GO; GO:0006807; P:nitrogen compound metabolism; IEA.
DR   InterPro; IPR002639; UreF.
DR   Pfam; PF01730; UreF; 1.
DR   PIRSF; PIRSF009467; Ureas_acces_UreF; 1.
SQ   SEQUENCE   228 AA;  25042 MW;  A3AE79A5B8C07092 CRC64;

  Query Match             70.4%;  Score 38;  DB 2;  Length 228;
  Best Local Similarity   77.8%;  Pred. No. 34;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GVREVIVMH 9
              | |||:|||
Db        147 GQREVVVMH 155


Search completed: March 31, 2006, 16:35:05
Job time : 43.209 secs
```

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: March 31, 2006, 16:09:06 ; Search time 53.9801 Seconds
(without alignments)
113.955 Million cell updates/sec

Title: US-10-609-217-83
Perfect score: 61
Sequence: 1 YXCXXGPXTWXCXP 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	58	95.1	20	2	AAY13728	Aay13728 Erythropo
2	58	95.1	20	2	AAY13687	Aay13687 Erythropo
3	58	95.1	20	2	AAW27001	Aaw27001 Monomer s
4	58	95.1	21	9	ADU91978	Adu91978 EPO-R ago
5	58	95.1	22	2	AAY13709	Aay13709 Erythropo
6	58	95.1	22	2	AAY26355	Aay26355 Erythropo
7	58	95.1	22	2	AAW27023	Aaw27023 Monomer s
8	57	93.4	17	9	ADU91963	Adu91963 EPO-R ago
9	57	93.4	17	9	ADU92005	Adu92005 EPO-R ago
10	57	93.4	20	2	AAY26409	Aay26409 Erythropo
11	57	93.4	20	2	AAY13650	Aay13650 Erythropo
12	57	93.4	20	2	AAY13676	Aay13676 Erythropo
13	57	93.4	20	2	AAY13630	Aay13630 Erythropo
14	57	93.4	20	2	AAY26383	Aay26383 Erythropo
15	57	93.4	20	2	AAW26990	Aaw26990 Monomer s
16	57	93.4	20	2	AAW26966	Aaw26966 Monomer s
17	57	93.4	20	2	AAW27042	Aaw27042 Monomer s
18	57	93.4	20	3	AAB17033	Aab17033 EPO-mimet
19	57	93.4	20	3	AAB17926	Aab17926 EPO-mimet
20	57	93.4	20	3	AAB13504	Aab13504 Erythropo
21	57	93.4	20	5	ABB72816	Abb72816 Erythropo
22	57	93.4	20	5	ABB72843	Abb72843 Erythropo
23	57	93.4	20	5	AAU74459	Aau74459 Human ery
24	57	93.4	20	5	AAU74480	Aau74480 Human ery
25	57	93.4	20	7	ADJ72555	Adj72555 EPO mimet
26	57	93.4	20	7	ADJ72582	Adj72582 EPO mimet
27	57	93.4	20	8	ADH10378	Adh10378 Erythropo
28	57	93.4	20	8	ADJ52191	Adj52191 CH1 delet
29	57	93.4	20	8	ADJ52218	Adj52218 CH1 delet
30	57	93.4	20	8	ADJ51180	Adj51180 CH1 delet
31	57	93.4	20	8	ADJ51153	Adj51153 CH1 delet
32	57	93.4	20	8	ADU47106	Adu47106 Erythropo
33	57	93.4	20	9	ADZ44405	Adz44405 Erythropo
34	57	93.4	20	9	ADZ44426	Adz44426 Erythropo
35	57	93.4	21	9	ADU91939	Adu91939 EPO-R ago
36	57	93.4	21	9	ADU91989	Adu91989 EPO-R ago
37	57	93.4	21	9	ADU91929	Adu91929 EPO-R ago
38	57	93.4	22	2	AAY13715	Aay13715 Erythropo
39	57	93.4	22	2	AAY26363	Aay26363 Erythropo
40	57	93.4	22	2	AAW27029	Aaw27029 Monomer s
41	57	93.4	26	2	AAY26445	Aay26445 Erythropo
42	57	93.4	26	2	AAY26441	Aay26441 Erythropo
43	57	93.4	133	7	ADJ73535	Adj73535 Erythropo
44	56	91.8	14	2	AAY13654	Aay13654 Erythropo
45	56	91.8	14	5	AAU74472	Aau74472 Human ery

ALIGNMENTS

RESULT 1
AAY13728
ID AAY13728 standard; peptide; 20 AA.
XX
AC AAY13728;
XX
DT 06-SEP-1999 (first entry)
XX
DE Erythropoietin receptor (EPO-R) binding peptide.
XX
KW Erythropoietin; EPO; receptor; EPO deficiency; renal failure; AIDS;
KW dialysis; anaemia; autoimmune disease; chronic inflammatory disease;
KW malignancy; beta-thalassaemia; cystic fibrosis; prematurity; blood loss;
KW spinal cord injury; aging; neoplastic disease; erythropoiesis; EPO-R.
XX
OS Synthetic.
XX
PN WO9640749-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US009810.
XX
PR 07-JUN-1995; 95US-00484631.
PR 07-JUN-1995; 95US-00484635.
XX
PA (JOHJ) JOHNSON & JOHNSON CORP.
PA (AFFY-) AFFYMAX TECHNOLOGIES NV.
XX
PI Wrighton NC, Dower WJ, Chang RS, Kashyap AK, Jolliffe LK;
PI Johnson D, Mulcahy L;
XX
DR WPI; 1997-052225/05.
XX
PT Erythropoietin receptor binding peptide - useful for treating disorders
PT characterised by deficiency of EPO, or low or defective red blood cell
PT population.
XX
PS Disclosure; Fig 2; 95pp; English.
XX
CC The invention describes a peptide of 10-40 amino acid residues which
CC binds to erythropoietin (EPO) receptor and which includes the amino acid
CC sequence Cys-Xaa1-Xaa2-Gly-Pro-Xaa3-Thr-Trp-Xaa4-Cys, where Xaa1 = Arg,
CC His, Leu or Trp, Xaa2 = Met, Phe or Ile, Xaa3 = 1 of the 20 genetically
CC coded L-amino acids, and Xaa4 = Asp, Glu, Ile, Leu or Val. Optionally,
CC the peptide may be cyclised or dimerised. The peptide can be used to
CC treat a patient having a disorder characterised by a deficiency of EPO or

CC a low or defective red blood cell population. It can be used to treat end
CC stage renal failure or dialysis; anaemia associated with AIDS, autoimmune
CC disease, chronic inflammatory diseases or malignancy; beta-thalassaemia;
CC cystic fibrosis; early anaemia of prematurity; spinal cord injury; acute
CC blood loss; aging; and neoplastic disease states accompanied by abnormal
CC erythropoiesis. The peptides can also be used as reagents for detecting
CC EPO receptors on living cells, in biological fluids, in tissue
CC homogenates, etc. Sequences AAY13662-735 are representative peptides of
CC the invention
XX
SQ Sequence 20 AA;

Query Match 95.1%; Score 58; DB 2; Length 20;
Best Local Similarity 57.1%; Pred. No. 0.015;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 YXCXXGPXTWXCXP 14
| | || || | |
Db 4 YACRMGPTTWVCSP 17

RESULT 2
AAY13687
ID AAY13687 standard; peptide; 20 AA.
XX
AC AAY13687;
XX
DT 06-SEP-1999 (first entry)
XX
DE Erythropoietin receptor (EPO-R) binding peptide.
XX
KW Erythropoietin; EPO; receptor; EPO deficiency; renal failure; AIDS;
KW dialysis; anaemia; autoimmune disease; chronic inflammatory disease;
KW malignancy; beta-thalassaemia; cystic fibrosis; prematurity; blood loss;
KW spinal cord injury; aging; neoplastic disease; erythropoiesis; EPO-R.
XX
OS Synthetic.
XX
PN WO9640749-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US009810.
XX
PR 07-JUN-1995; 95US-00484631.
PR 07-JUN-1995; 95US-00484635.
XX
PA (JOHJ) JOHNSON & JOHNSON CORP.
PA (AFFY-) AFFYMAX TECHNOLOGIES NV.
XX
PI Wrighton NC, Dower WJ, Chang RS, Kashyap AK, Jolliffe LK;
PI Johnson D, Mulcahy L;
XX
DR WPI; 1997-052225/05.
XX
PT Erythropoietin receptor binding peptide - useful for treating disorders
PT characterised by deficiency of EPO, or low or defective red blood cell
PT population.
XX
PS Disclosure; Fig 2; 95pp; English.
XX
CC The invention describes a peptide of 10-40 amino acid residues which
CC binds to erythropoietin (EPO) receptor and which includes the amino acid
CC sequence Cys-Xaa1-Xaa2-Gly-Pro-Xaa3-Thr-Trp-Xaa4-Cys, where Xaa1 = Arg,
CC His, Leu or Trp, Xaa2 = Met, Phe or Ile, Xaa3 = 1 of the 20 genetically
CC coded L-amino acids, and Xaa4 = Asp, Glu, Ile, Leu or Val. Optionally,
CC the peptide may be cyclised or dimerised. The peptide can be used to
CC treat a patient having a disorder characterised by a deficiency of EPO or
CC a low or defective red blood cell population. It can be used to treat end
CC stage renal failure or dialysis; anaemia associated with AIDS, autoimmune
CC disease, chronic inflammatory diseases or malignancy; beta-thalassaemia;
CC cystic fibrosis; early anaemia of prematurity; spinal cord injury; acute
CC blood loss; aging; and neoplastic disease states accompanied by abnormal
CC erythropoiesis. The peptides can also be used as reagents for detecting
CC EPO receptors on living cells, in biological fluids, in tissue
CC homogenates, etc. Sequences AAY13662-735 are representative peptides of
CC the invention
XX
SQ Sequence 20 AA;

Query Match 95.1%; Score 58; DB 2; Length 20;
Best Local Similarity 57.1%; Pred. No. 0.015;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 YXCXXGPXTWXCXP 14
| | || || | |
Db 4 YSCRMGPTTWVCTP 17

RESULT 3
AAW27001
ID AAW27001 standard; peptide; 20 AA.
XX
AC AAW27001;
XX
DT 11-NOV-1997 (first entry)
XX
DE Monomer subunit of erythropoietin receptor binding dimer.
XX
KW Monomer; erythropoietin; EPO; receptor; binding; dimer; activation;
KW treatment; disorder; deficiency; low; defective; red blood cell;
KW erythrocyte; population; cell surface; agonist; end stage; renal;
KW failure; dialysis; anaemia; anemia; AIDS; chronic; inflammatory; disease;
KW rheumatoid arthritis; bowel inflammation; autoimmune; transfusion.
XX
OS Synthetic.
XX
PN WO9640772-A2.
XX
PD 19-DEC-1996.
XX
PF 06-JUN-1996; 96WO-US009469.
XX
PR 07-JUN-1995; 95US-00484135.
XX
PA (JOHJ) JOHNSON & JOHNSON.
XX
PI Johnson DL, Zivin RA;
XX
DR WPI; 1997-099920/09.
XX
PT Activating cell surface receptors using peptide dimer agonists - also,
PT new dimers of erythropoietin receptor binding peptide(s) useful for
PT treating patient having disorder characterised by EPO deficiency.
XX
PS Disclosure; Fig 9; 110pp; English.
XX
CC The present peptide is a specific example of a claimed generic monomer
CC subunit of an erythropoietin (EPO) receptor binding dimer, which
CC comprises 2 EPO receptor binding monomers of 10 to 40 amino acids, and
CC activates or improves the bioactivity of the EPO cell surface receptor.
CC The dimer can be used to treat disorders resulting from EPO deficiency by
CC improving the activity of its cell surface receptor, e.g. end stage renal
CC failure/dialysis, anaemia associated with AIDS or chronic inflammatory
CC diseases such as rheumatoid arthritis and chronic bowel inflammation and
CC autoimmune disease. It can also be used to boost the red cell count of a
CC patient prior to surgery or as pretreatment to transfusion. The dimer
CC peptide exhibits increased biological potency in vitro and in vivo
CC relative to its component monomeric agonists. Dimerisation may also
CC convert cell surface receptor antagonists into agonists
XX
SQ Sequence 20 AA;

Query Match 95.1%; Score 58; DB 2; Length 20;
Best Local Similarity 57.1%; Pred. No. 0.015;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
Qy          1 YXCXXGPXTWXCXP 14
              | |  || || | |
Db          4 YSCRMGPTTWVCTP 17
```

RESULT 4
ADU91978
ID ADU91978 standard; peptide; 21 AA.
XX
AC ADU91978;
XX
DT 10-FEB-2005 (first entry)
XX
DE EPO-R agonist SEQ ID NO 119.
XX
KW erythropoietin receptor; EPO-R; erythropoietin; renal failure;
KW autoimmune disease; cystic fibrosis; anemia; inflammation;
KW spinal cord injury; aging; neurological disease; nephrotropic;
KW antianemic; immunosuppressive; CNS-Gen.; neuroprotective;
KW respiratory-Gen.; antiinflammatory; vulnerary; nootropic; cytostatic;
KW hemostatic; cyclic.
XX
OS Synthetic.
XX
FH Key   Location/Qualifiers
FT Modified-site   1
FT   /note= "Acetylated residue"
FT Disulfide-bond 7. .16
FT Modified-site   21
FT   /note= "C-terminal amide"
XX
PN WO2004101611-A2.
XX
PD 25-NOV-2004.
XX
PF 12-MAY-2004; 2004WO-US014886.
XX
PR 12-MAY-2003; 2003US-0470245P.
XX
PA (AFFY-) AFFYMAX INC.
XX
PI Yin K, Holmes C, Lalonde G, Balu P, Schatz PJ, Tumelty D;
XX
DR WPI; 2005-039329/04.
XX
PT New peptide comprising specified sequence of amino acid is erythropoietin
PT receptor agonist useful for treating e.g. anemia, beta-thalassemia, renal
PT disorders.
XX
PS Disclosure; SEQ ID NO 119; 83pp; English.
XX
CC This invention describes a novel peptide which is an erythropoietin
CC receptor (EPO-R) activator. The peptide forms a dimer comprising a
CC linking moiety connecting two peptide chains composed of ADU91861. The N-
CC terminal of the peptide is acetylated. The EPO-R activator further
CC comprises at least one water soluble polymer, preferably polyethylene
CC glycol (PEG) covalently bound to the peptide and a spacer moiety. The
CC products of the invention are used for treating disorders associated with
CC deficiency of erythropoietin or low or defective red blood cell
CC population, end stage renal failure or dialysis, anemia associated with
CC AIDS, autoimmune disease or malignancy, beta-thalassemia, cystic
CC fibrosis, early anemia of prematurity, anemia associated with chronic
CC inflammatory disease, spinal cord injury, acute blood loss, aging and
CC neoplastic disease states accompanied by abnormal erythropoiesis. The
CC peptide compounds are potent agonists of erythropoietin receptor and have
CC nephrotropic, antianemic, immunosuppressive, CNS-Gen., neuroprotective,
CC respiratory-Gen., antiinflammatory, vulnerary, nootropic, cytostatic and
CC hemostatic activity. This sequence represents a peptide which acts as an
CC erythropoietin receptor (EPO-R) agonist.
XX
SQ Sequence 21 AA;

```
  Query Match             95.1%;  Score 58;  DB 9;  Length 21;
  Best Local Similarity   57.1%;  Pred. No. 0.016;
  Matches    8;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 YXCXXGPXTWXCXP 14
              | |  || || | |
Db          5 YSCQMGPTTWTCVP 18
```

RESULT 5
AAY13709
ID AAY13709 standard; peptide; 22 AA.
XX
AC AAY13709;
XX
DT 06-SEP-1999 (first entry)
XX
DE Erythropoietin receptor (EPO-R) binding peptide.
XX
KW Erythropoietin; EPO; receptor; EPO deficiency; renal failure; AIDS;
KW dialysis; anaemia; autoimmune disease; chronic inflammatory disease;
KW malignancy; beta-thalassaemia; cystic fibrosis; prematurity; blood loss;
KW spinal cord injury; aging; neoplastic disease; erythropoiesis; EPO-R.
XX
OS Synthetic.
XX
PN WO9640749-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US009810.
XX
PR 07-JUN-1995; 95US-00484631.
PR 07-JUN-1995; 95US-00484635.
XX
PA (JOHJ) JOHNSON & JOHNSON CORP.
PA (AFFY-) AFFYMAX TECHNOLOGIES NV.
XX
PI Wrighton NC, Dower WJ, Chang RS, Kashyap AK, Jolliffe LK;
PI Johnson D, Mulcahy L;
XX
DR WPI; 1997-052225/05.
XX
PT Erythropoietin receptor binding peptide - useful for treating disorders
PT characterised by deficiency of EPO, or low or defective red blood cell
PT population.
XX
PS Disclosure; Fig 2; 95pp; English.
XX
CC The invention describes a peptide of 10-40 amino acid residues which
CC binds to erythropoietin (EPO) receptor and which includes the amino acid
CC sequence Cys-Xaa1-Xaa2-Gly-Pro-Xaa3-Thr-Trp-Xaa4-Cys, where Xaa1 = Arg,
CC His, Leu or Trp, Xaa2 = Met, Phe or Ile, Xaa3 = 1 of the 20 genetically
CC coded L-amino acids, and Xaa4 = Asp, Glu, Ile, Leu or Val. Optionally,
CC the peptide may be cyclised or dimerised. The peptide can be used to
CC treat a patient having a disorder characterised by a deficiency of EPO or
CC a low or defective red blood cell population. It can be used to treat end
CC stage renal failure or dialysis; anaemia associated with AIDS, autoimmune
CC disease, chronic inflammatory diseases or malignancy; beta-thalassaemia;
CC cystic fibrosis; early anaemia of prematurity; spinal cord injury; acute
CC blood loss; aging; and neoplastic disease states accompanied by abnormal
CC erythropoiesis. The peptides can also be used as reagents for detecting
CC EPO receptors on living cells, in biological fluids, in tissue
CC homogenates, etc. Sequences AAY13662-735 are representative peptides of
CC the invention
XX
SQ Sequence 22 AA;

```
  Query Match             95.1%;  Score 58;  DB 2;  Length 22;
  Best Local Similarity   57.1%;  Pred. No. 0.017;
  Matches    8;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 YXCXXGPXTWXCXP 14
```

```
              | |  || || | |
Db          4 YSCFMGPTTWVCSP 17


RESULT 6
AAY26355
ID   AAY26355 standard; peptide; 22 AA.
XX
AC   AAY26355;
XX
DT   06-SEP-1999  (first entry)
XX
DE   Erythropoietin receptor (EPO-R) binding peptide.
XX
KW   Erythropoietin; EPO; receptor; EPO deficiency; renal failure; AIDS;
KW   dialysis; anaemia; autoimmune disease; chronic inflammatory disease;
KW   malignancy; beta-thalassaemia; cystic fibrosis; prematurity; blood loss;
KW   spinal cord injury; aging; neoplastic disease; erythropoiesis; EPO-R.
XX
OS   Synthetic.
XX
PN   WO9640749-A1.
XX
PD   19-DEC-1996.
XX
PF   07-JUN-1996;   96WO-US009810.
XX
PR   07-JUN-1995;   95US-00484631.
PR   07-JUN-1995;   95US-00484635.
XX
PA   (JOHJ ) JOHNSON & JOHNSON CORP.
PA   (AFFY-) AFFYMAX TECHNOLOGIES NV.
XX
PI   Wrighton NC,  Dower WJ,  Chang RS,  Kashyap AK,  Jolliffe LK;
PI   Johnson D,  Mulcahy L;
XX
DR   WPI; 1997-052225/05.
XX
PT   Erythropoietin receptor binding peptide - useful for treating disorders
PT   characterised by deficiency of EPO, or low or defective red blood cell
PT   population.
XX
PS   Disclosure; Page 16; 95pp; English.
XX
CC   The invention describes a peptide of 10-40 amino acid residues which
CC   binds to erythropoietin (EPO) receptor and which includes the amino acid
CC   sequence Cys-Xaa1-Xaa2-Gly-Pro-Xaa3-Thr-Trp-Xaa4-Cys, where Xaa1 = Arg,
CC   His, Leu or Trp, Xaa2 = Met, Phe or Ile, Xaa3 = 1 of the 20 genetically
CC   coded L-amino acids, and Xaa4 = Asp, Glu, Ile, Leu or Val. Optionally,
CC   the peptide may be cyclised or dimerised. The peptide can be used to
CC   treat a patient having a disorder characterised by a deficiency of EPO or
CC   a low or defective red blood cell population. It can be used to treat end
CC   stage renal failure or dialysis; anaemia associated with AIDS, autoimmune
CC   disease, chronic inflammatory diseases or malignancy; beta-thalassaemia;
CC   cystic fibrosis; early anaemia of prematurity; spinal cord injury; acute
CC   blood loss; aging; and neoplastic disease states accompanied by abnormal
CC   erythropoiesis. The peptides can also be used as reagents for detecting
CC   EPO receptors on living cells, in biological fluids, in tissue
CC   homogenates, etc. Sequences AAY26352-548 are representative peptides
CC   falling within the above peptide motif and isolated during the affinity
CC   selection process
XX
SQ   Sequence 22 AA;

  Query Match             95.1%;  Score 58;  DB 2;  Length 22;
  Best Local Similarity   57.1%;  Pred. No. 0.017;
  Matches    8;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 YXCXXGPXTWXCXP 14
              | |  || || | |
Db          4 YSCFMGPTTWVCSP 17


RESULT 7
AAW27023
ID   AAW27023 standard; peptide; 22 AA.
XX
AC   AAW27023;
XX
DT   11-NOV-1997  (first entry)
XX
DE   Monomer subunit of erythropoietin receptor binding dimer.
XX
KW   Monomer; erythropoietin; EPO; receptor; binding; dimer; activation;
KW   treatment; disorder; deficiency; low; defective; red blood cell;
KW   erythrocyte; population; cell surface; agonist; end stage; renal;
KW   failure; dialysis; anaemia; anemia; AIDS; chronic; inflammatory; disease;
KW   rheumatoid arthritis; bowel inflammation; autoimmune; transfusion.
XX
OS   Synthetic.
XX
PN   WO9640772-A2.
XX
PD   19-DEC-1996.
XX
PF   06-JUN-1996;   96WO-US009469.
XX
PR   07-JUN-1995;   95US-00484135.
XX
PA   (JOHJ ) JOHNSON & JOHNSON.
XX
PI   Johnson DL,  Zivin RA;
XX
DR   WPI; 1997-099920/09.
XX
PT   Activating cell surface receptors using peptide dimer agonists - also,
PT   new dimers of erythropoietin receptor binding peptide(s) useful for
PT   treating patient having disorder characterised by EPO deficiency.
XX
PS   Disclosure; Fig 9; 110pp; English.
XX
CC   The present peptide is a specific example of a claimed generic monomer
CC   subunit of an erythropoietin (EPO) receptor binding dimer, which
CC   comprises 2 EPO receptor binding monomers of 10 to 40 amino acids, and
CC   activates or improves the bioactivity of the EPO cell surface receptor.
CC   The dimer can be used to treat disorders resulting from EPO deficiency by
CC   improving the activity of its cell surface receptor, e.g. end stage renal
CC   failure/dialysis, anaemia associated with AIDS or chronic inflammatory
CC   diseases such as rheumatoid arthritis and chronic bowel inflammation and
CC   autoimmune disease. It can also be used to boost the red cell count of a
CC   patient prior to surgery or as pretreatment to transfusion. The dimer
CC   peptide exhibits increased biological potency in vitro and in vivo
CC   relative to its component monomeric agonists. Dimerisation may also
CC   convert cell surface receptor antagonists into agonists
XX
SQ   Sequence 22 AA;

  Query Match             95.1%;  Score 58;  DB 2;  Length 22;
  Best Local Similarity   57.1%;  Pred. No. 0.017;
  Matches    8;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 YXCXXGPXTWXCXP 14
              | |  || || | |
Db          4 YSCFMGPTTWVCSP 17


RESULT 8
ADU91963
ID   ADU91963 standard; peptide; 17 AA.
XX
AC   ADU91963;
XX
DT   10-FEB-2005  (first entry)
XX
DE   EPO-R agonist SEQ ID NO 104.
XX
```

```
KW   erythropoietin receptor; EPO-R; erythropoietin; renal failure;
KW   autoimmune disease; cystic fibrosis; anemia; inflammation;
KW   spinal cord injury; aging; neurological disease; nephrotropic;
KW   antianemic; immunosuppressive; CNS-Gen.; neuroprotective;
KW   respiratory-Gen.; antiinflammatory; vulnerary; nootropic; cytostatic;
KW   hemostatic; cyclic.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "Acetylated residue"
FT   Disulfide-bond  4. .13
FT   Modified-site   17
FT                   /note= "C-terminal amide"
XX
PN   WO2004101611-A2.
XX
PD   25-NOV-2004.
XX
PF   12-MAY-2004; 2004WO-US014886.
XX
PR   12-MAY-2003; 2003US-0470245P.
XX
PA   (AFFY-) AFFYMAX INC.
XX
PI   Yin K,  Holmes C,  Lalonde G,  Balu P,  Schatz PJ,  Tumelty D;
XX
DR   WPI; 2005-039329/04.
XX
PT   New peptide comprising specified sequence of amino acid is erythropoietin
PT   receptor agonist useful for treating e.g. anemia, beta-thalassemia, renal
PT   disorders.
XX
PS   Disclosure; SEQ ID NO 104; 83pp; English.
XX
CC   This invention describes a novel peptide which is an erythropoietin
CC   receptor (EPO-R) activator. The peptide forms a dimer comprising a
CC   linking moiety connecting two peptide chains composed of ADU91861. The N-
CC   terminal of the peptide is acetylated. The EPO-R activator further
CC   comprises at least one water soluble polymer, preferably polyethylene
CC   glycol (PEG) covalently bound to the peptide and a spacer moiety. The
CC   products of the invention are used for treating disorders associated with
CC   deficiency of erythropoietin or low or defective red blood cell
CC   population, end stage renal failure or dialysis, anemia associated with
CC   AIDS, autoimmune disease or malignancy, beta-thalassemia, cystic
CC   fibrosis, early anemia of prematurity, anemia associated with chronic
CC   inflammatory disease, spinal cord injury, acute blood loss, aging and
CC   neoplastic disease states accompanied by abnormal erythropoiesis. The
CC   peptide compounds are potent agonists of erythropoietin receptor and have
CC   nephrotropic, antianemic, immunosuppressive, CNS-Gen., neuroprotective,
CC   respiratory-Gen., antiinflammatory, vulnerary, nootropic, cytostatic and
CC   hemostatic activity. This sequence represents a peptide which acts as an
CC   erythropoietin receptor (EPO-R) agonist.
XX
SQ   Sequence 17 AA;

  Query Match             93.4%;  Score 57;  DB 9;  Length 17;
  Best Local Similarity   57.1%;  Pred. No. 0.019;
  Matches    8;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 YXCXXGPXTWXCXP 14
              | |  || || | |
Db          2 YSCRMGPMTWVCSP 15


RESULT 9
ADU92005
ID   ADU92005 standard; peptide; 17 AA.
XX
AC   ADU92005;
XX
DT   10-FEB-2005  (first entry)
XX
DE   EPO-R agonist SEQ ID NO 146.
XX
KW   erythropoietin receptor; EPO-R; erythropoietin; renal failure;
KW   autoimmune disease; cystic fibrosis; anemia; inflammation;
KW   spinal cord injury; aging; neurological disease; nephrotropic;
KW   antianemic; immunosuppressive; CNS-Gen.; neuroprotective;
KW   respiratory-Gen.; antiinflammatory; vulnerary; nootropic; cytostatic;
KW   hemostatic; cyclic.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "Acetylated residue"
FT   Disulfide-bond  4. .13
FT   Modified-site   17
FT                   /note= "C-terminal amide"
XX
PN   WO2004101611-A2.
XX
PD   25-NOV-2004.
XX
PF   12-MAY-2004; 2004WO-US014886.
XX
PR   12-MAY-2003; 2003US-0470245P.
XX
PA   (AFFY-) AFFYMAX INC.
XX
PI   Yin K,  Holmes C,  Lalonde G,  Balu P,  Schatz PJ,  Tumelty D;
XX
DR   WPI; 2005-039329/04.
XX
PT   New peptide comprising specified sequence of amino acid is erythropoietin
PT   receptor agonist useful for treating e.g. anemia, beta-thalassemia, renal
PT   disorders.
XX
PS   Disclosure; SEQ ID NO 146; 83pp; English.
XX
CC   This invention describes a novel peptide which is an erythropoietin
CC   receptor (EPO-R) activator. The peptide forms a dimer comprising a
CC   linking moiety connecting two peptide chains composed of ADU91861. The N-
CC   terminal of the peptide is acetylated. The EPO-R activator further
CC   comprises at least one water soluble polymer, preferably polyethylene
CC   glycol (PEG) covalently bound to the peptide and a spacer moiety. The
CC   products of the invention are used for treating disorders associated with
CC   deficiency of erythropoietin or low or defective red blood cell
CC   population, end stage renal failure or dialysis, anemia associated with
CC   AIDS, autoimmune disease or malignancy, beta-thalassemia, cystic
CC   fibrosis, early anemia of prematurity, anemia associated with chronic
CC   inflammatory disease, spinal cord injury, acute blood loss, aging and
CC   neoplastic disease states accompanied by abnormal erythropoiesis. The
CC   peptide compounds are potent agonists of erythropoietin receptor and have
CC   nephrotropic, antianemic, immunosuppressive, CNS-Gen., neuroprotective,
CC   respiratory-Gen., antiinflammatory, vulnerary, nootropic, cytostatic and
CC   hemostatic activity. This sequence represents a peptide which acts as an
CC   erythropoietin receptor (EPO-R) agonist.
XX
SQ   Sequence 17 AA;

  Query Match             93.4%;  Score 57;  DB 9;  Length 17;
  Best Local Similarity   57.1%;  Pred. No. 0.019;
  Matches    8;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 YXCXXGPXTWXCXP 14
              | |  || || | |
Db          2 YTCRFGPLTWECTP 15


RESULT 10
AAY26409
ID   AAY26409 standard; peptide; 20 AA.
XX
```

```
AC   AAY26409;
XX
DT   06-SEP-1999  (first entry)
XX
DE   Erythropoietin receptor (EPO-R) binding peptide.
XX
KW   Erythropoietin; EPO; receptor; EPO deficiency; renal failure; AIDS;
KW   dialysis; anaemia; autoimmune disease; chronic inflammatory disease;
KW   malignancy; beta-thalassaemia; cystic fibrosis; prematurity; blood loss;
KW   spinal cord injury; aging; neoplastic disease; erythropoiesis; EPO-R.
XX
OS   Synthetic.
XX
PN   WO9640749-A1.
XX
PD   19-DEC-1996.
XX
PF   07-JUN-1996;   96WO-US009810.
XX
PR   07-JUN-1995;   95US-00484631.
PR   07-JUN-1995;   95US-00484635.
XX
PA   (JOHJ ) JOHNSON & JOHNSON CORP.
PA   (AFFY-) AFFYMAX TECHNOLOGIES NV.
XX
PI   Wrighton NC,  Dower WJ,  Chang RS,  Kashyap AK,  Jolliffe LK;
PI   Johnson D,  Mulcahy L;
XX
DR   WPI; 1997-052225/05.
XX
PT   Erythropoietin receptor binding peptide - useful for treating disorders
PT   characterised by deficiency of EPO, or low or defective red blood cell
PT   population.
XX
PS   Disclosure; Page 19; 95pp; English.
XX
CC   The invention describes a peptide of 10-40 amino acid residues which
CC   binds to erythropoietin (EPO) receptor and which includes the amino acid
CC   sequence Cys-Xaa1-Xaa2-Gly-Pro-Xaa3-Thr-Trp-Xaa4-Cys, where Xaa1 = Arg,
CC   His, Leu or Trp, Xaa2 = Met, Phe or Ile, Xaa3 = 1 of the 20 genetically
CC   coded L-amino acids, and Xaa4 = Asp, Glu, Ile, Leu or Val. Optionally,
CC   the peptide may be cyclised or dimerised. The peptide can be used to
CC   treat a patient having a disorder characterised by a deficiency of EPO or
CC   a low or defective red blood cell population. It can be used to treat end
CC   stage renal failure or dialysis; anaemia associated with AIDS, autoimmune
CC   disease, chronic inflammatory diseases or malignancy; beta-thalassaemia;
CC   cystic fibrosis; early anaemia of prematurity; spinal cord injury; acute
CC   blood loss; aging; and neoplastic disease states accompanied by abnormal
CC   erythropoiesis. The peptides can also be used as reagents for detecting
CC   EPO receptors on living cells, in biological fluids, in tissue
CC   homogenates, etc. Sequences AAY26352-548 are representative peptides
CC   falling within the above peptide motif and isolated during the affinity
CC   selection process
XX
SQ   Sequence 20 AA;

  Query Match             93.4%;  Score 57;  DB 2;  Length 20;
  Best Local Similarity   57.1%;  Pred. No. 0.022;
  Matches    8;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 YXCXXGPXTWXCXP 14
              | |  || || | |
Db          4 YACRMGPITWVCSP 17


RESULT 11
AAY13650
ID   AAY13650 standard; peptide; 20 AA.
XX
AC   AAY13650;
XX
DT   06-SEP-1999  (first entry)
XX
DE   Erythropoietin receptor (EPO-R) binding peptide.
XX
KW   Erythropoietin; EPO; receptor; EPO deficiency; renal failure; AIDS;
KW   dialysis; anaemia; autoimmune disease; chronic inflammatory disease;
KW   malignancy; beta-thalassaemia; cystic fibrosis; prematurity; blood loss;
KW   spinal cord injury; aging; neoplastic disease; erythropoiesis; EPO-R.
XX
OS   Synthetic.
XX
PN   WO9640749-A1.
XX
PD   19-DEC-1996.
XX
PF   07-JUN-1996;   96WO-US009810.
XX
PR   07-JUN-1995;   95US-00484631.
PR   07-JUN-1995;   95US-00484635.
XX
PA   (JOHJ ) JOHNSON & JOHNSON CORP.
PA   (AFFY-) AFFYMAX TECHNOLOGIES NV.
XX
PI   Wrighton NC,  Dower WJ,  Chang RS,  Kashyap AK,  Jolliffe LK;
PI   Johnson D,  Mulcahy L;
XX
DR   WPI; 1997-052225/05.
XX
PT   Erythropoietin receptor binding peptide - useful for treating disorders
PT   characterised by deficiency of EPO, or low or defective red blood cell
PT   population.
XX
PS   Claim 6; Page 68; 95pp; English.
XX
CC   The invention describes a peptide of 10-40 amino acid residues which
CC   binds to erythropoietin (EPO) receptor and which includes the amino acid
CC   sequence Cys-Xaa1-Xaa2-Gly-Pro-Xaa3-Thr-Trp-Xaa4-Cys, where Xaa1 = Arg,
CC   His, Leu or Trp, Xaa2 = Met, Phe or Ile, Xaa3 = 1 of the 20 genetically
CC   coded L-amino acids, and Xaa4 = Asp, Glu, Ile, Leu or Val. Optionally,
CC   the peptide may be cyclised or dimerised. The peptide can be used to
CC   treat a patient having a disorder characterised by a deficiency of EPO or
CC   a low or defective red blood cell population. It can be used to treat end
CC   stage renal failure or dialysis; anaemia associated with AIDS, autoimmune
CC   disease, chronic inflammatory diseases or malignancy; beta-thalassaemia;
CC   cystic fibrosis; early anaemia of prematurity; spinal cord injury; acute
CC   blood loss; aging; and neoplastic disease states accompanied by abnormal
CC   erythropoiesis. The peptides can also be used as reagents for detecting
CC   EPO receptors on living cells, in biological fluids, in tissue
CC   homogenates, etc. Sequences AAY13624-661 represent specifc examples of
CC   EPO-R binding peptides
XX
SQ   Sequence 20 AA;

  Query Match             93.4%;  Score 57;  DB 2;  Length 20;
  Best Local Similarity   57.1%;  Pred. No. 0.022;
  Matches    8;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 YXCXXGPXTWXCXP 14
              | |  || || | |
Db          4 YSCHFGPATWVCKP 17


RESULT 12
AAY13676
ID   AAY13676 standard; peptide; 20 AA.
XX
AC   AAY13676;
XX
DT   06-SEP-1999  (first entry)
XX
DE   Erythropoietin receptor (EPO-R) binding peptide.
XX
KW   Erythropoietin; EPO; receptor; EPO deficiency; renal failure; AIDS;
KW   dialysis; anaemia; autoimmune disease; chronic inflammatory disease;
KW   malignancy; beta-thalassaemia; cystic fibrosis; prematurity; blood loss;
```

KW spinal cord injury; aging; neoplastic disease; erythropoiesis; EPO-R.
XX
OS Synthetic.
XX
PN WO9640749-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US009810.
XX
PR 07-JUN-1995; 95US-00484631.
PR 07-JUN-1995; 95US-00484635.
XX
PA (JOHJ) JOHNSON & JOHNSON CORP.
PA (AFFY-) AFFYMAX TECHNOLOGIES NV.
XX
PI Wrighton NC, Dower WJ, Chang RS, Kashyap AK, Jolliffe LK;
PI Johnson D, Mulcahy L;
XX
DR WPI; 1997-052225/05.
XX
PT Erythropoietin receptor binding peptide - useful for treating disorders
PT characterised by deficiency of EPO, or low or defective red blood cell
PT population.
XX
PS Disclosure; Fig 2; 95pp; English.
XX
CC The invention describes a peptide of 10-40 amino acid residues which
CC binds to erythropoietin (EPO) receptor and which includes the amino acid
CC sequence Cys-Xaa1-Xaa2-Gly-Pro-Xaa3-Thr-Trp-Xaa4-Cys, where Xaa1 = Arg,
CC His, Leu or Trp, Xaa2 = Met, Phe or Ile, Xaa3 = 1 of the 20 genetically
CC coded L-amino acids, and Xaa4 = Asp, Glu, Ile, Leu or Val. Optionally,
CC the peptide may be cyclised or dimerised. The peptide can be used to
CC treat a patient having a disorder characterised by a deficiency of EPO or
CC a low or defective red blood cell population. It can be used to treat end
CC stage renal failure or dialysis; anaemia associated with AIDS, autoimmune
CC disease, chronic inflammatory diseases or malignancy; beta-thalassaemia;
CC cystic fibrosis; early anaemia of prematurity; spinal cord injury; acute
CC blood loss; aging; and neoplastic disease states accompanied by abnormal
CC erythropoiesis. The peptides can also be used as reagents for detecting
CC EPO receptors on living cells, in biological fluids, in tissue
CC homogenates, etc. Sequences AAY13662-735 are representative peptides of
CC the invention
XX
SQ Sequence 20 AA;

Query Match 93.4%; Score 57; DB 2; Length 20;
Best Local Similarity 57.1%; Pred. No. 0.022;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
Qy          1 YXCXXGPXTWXCXP 14
              | |  || || | |
Db          4 YSCRMGPMTWVCSP 17
```

RESULT 13
AAY13630
ID AAY13630 standard; peptide; 20 AA.
XX
AC AAY13630;
XX
DT 06-SEP-1999 (first entry)
XX
DE Erythropoietin receptor (EPO-R) binding peptide.
XX
KW Erythropoietin; EPO; receptor; EPO deficiency; renal failure; AIDS;
KW dialysis; anaemia; autoimmune disease; chronic inflammatory disease;
KW malignancy; beta-thalassaemia; cystic fibrosis; prematurity; blood loss;
KW spinal cord injury; aging; neoplastic disease; erythropoiesis; EPO-R.
XX
OS Synthetic.
XX
PN WO9640749-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US009810.
XX
PR 07-JUN-1995; 95US-00484631.
PR 07-JUN-1995; 95US-00484635.
XX
PA (JOHJ) JOHNSON & JOHNSON CORP.
PA (AFFY-) AFFYMAX TECHNOLOGIES NV.
XX
PI Wrighton NC, Dower WJ, Chang RS, Kashyap AK, Jolliffe LK;
PI Johnson D, Mulcahy L;
XX
DR WPI; 1997-052225/05.
XX
PT Erythropoietin receptor binding peptide - useful for treating disorders
PT characterised by deficiency of EPO, or low or defective red blood cell
PT population.
XX
PS Claim 6; Page 68; 95pp; English.
XX
CC The invention describes a peptide of 10-40 amino acid residues which
CC binds to erythropoietin (EPO) receptor and which includes the amino acid
CC sequence Cys-Xaa1-Xaa2-Gly-Pro-Xaa3-Thr-Trp-Xaa4-Cys, where Xaa1 = Arg,
CC His, Leu or Trp, Xaa2 = Met, Phe or Ile, Xaa3 = 1 of the 20 genetically
CC coded L-amino acids, and Xaa4 = Asp, Glu, Ile, Leu or Val. Optionally,
CC the peptide may be cyclised or dimerised. The peptide can be used to
CC treat a patient having a disorder characterised by a deficiency of EPO or
CC a low or defective red blood cell population. It can be used to treat end
CC stage renal failure or dialysis; anaemia associated with AIDS, autoimmune
CC disease, chronic inflammatory diseases or malignancy; beta-thalassaemia;
CC cystic fibrosis; early anaemia of prematurity; spinal cord injury; acute
CC blood loss; aging; and neoplastic disease states accompanied by abnormal
CC erythropoiesis. The peptides can also be used as reagents for detecting
CC EPO receptors on living cells, in biological fluids, in tissue
CC homogenates, etc. Sequences AAY13624-661 represent specifc examples of
CC EPO-R binding peptides
XX
SQ Sequence 20 AA;

Query Match 93.4%; Score 57; DB 2; Length 20;
Best Local Similarity 57.1%; Pred. No. 0.022;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
Qy          1 YXCXXGPXTWXCXP 14
              | |  || || | |
Db          4 YACRMGPITWVCSP 17
```

RESULT 14
AAY26383
ID AAY26383 standard; peptide; 20 AA.
XX
AC AAY26383;
XX
DT 06-SEP-1999 (first entry)
XX
DE Erythropoietin receptor (EPO-R) binding peptide.
XX
KW Erythropoietin; EPO; receptor; EPO deficiency; renal failure; AIDS;
KW dialysis; anaemia; autoimmune disease; chronic inflammatory disease;
KW malignancy; beta-thalassaemia; cystic fibrosis; prematurity; blood loss;
KW spinal cord injury; aging; neoplastic disease; erythropoiesis; EPO-R.
XX
OS Synthetic.
XX
PN WO9640749-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US009810.
XX

PR 07-JUN-1995; 95US-00484631.
PR 07-JUN-1995; 95US-00484635.
XX
PA (JOHJ) JOHNSON & JOHNSON CORP.
PA (AFFY-) AFFYMAX TECHNOLOGIES NV.
XX
PI Wrighton NC, Dower WJ, Chang RS, Kashyap AK, Jolliffe LK;
PI Johnson D, Mulcahy L;
XX
DR WPI; 1997-052225/05.
XX
PT Erythropoietin receptor binding peptide - useful for treating disorders
PT characterised by deficiency of EPO, or low or defective red blood cell
PT population.
XX
PS Disclosure; Page 17; 95pp; English.
XX
CC The invention describes a peptide of 10-40 amino acid residues which
CC binds to erythropoietin (EPO) receptor and which includes the amino acid
CC sequence Cys-Xaa1-Xaa2-Gly-Pro-Xaa3-Thr-Trp-Xaa4-Cys, where Xaa1 = Arg,
CC His, Leu or Trp, Xaa2 = Met, Phe or Ile, Xaa3 = 1 of the 20 genetically
CC coded L-amino acids, and Xaa4 = Asp, Glu, Ile, Leu or Val. Optionally,
CC the peptide may be cyclised or dimerised. The peptide can be used to
CC treat a patient having a disorder characterised by a deficiency of EPO or
CC a low or defective red blood cell population. It can be used to treat end
CC stage renal failure or dialysis; anaemia associated with AIDS, autoimmune
CC disease, chronic inflammatory diseases or malignancy; beta-thalassaemia;
CC cystic fibrosis; early anaemia of prematurity; spinal cord injury; acute
CC blood loss; aging; and neoplastic disease states accompanied by abnormal
CC erythropoiesis. The peptides can also be used as reagents for detecting
CC EPO receptors on living cells, in biological fluids, in tissue
CC homogenates, etc. Sequences AAY26352-548 are representative peptides
CC falling within the above peptide motif and isolated during the affinity
CC selection process
XX
SQ Sequence 20 AA;

Query Match 93.4%; Score 57; DB 2; Length 20;
Best Local Similarity 57.1%; Pred. No. 0.022;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 YXCXXGPXTWXCXP 14
|| || || | |
Db 4 YACRMGPITWVCSP 17

RESULT 15
AAW26990
ID AAW26990 standard; peptide; 20 AA.
XX
AC AAW26990;
XX
DT 11-NOV-1997 (first entry)
XX
DE Monomer subunit of erythropoietin receptor binding dimer.
XX
KW Monomer; erythropoietin; EPO; receptor; binding; dimer; activation;
KW treatment; disorder; deficiency; low; defective; red blood cell;
KW erythrocyte; population; cell surface; agonist; end stage; renal;
KW failure; dialysis; anaemia; anemia; AIDS; chronic; inflammatory; disease;
KW rheumatoid arthritis; bowel inflammation; autoimmune; transfusion.
XX
OS Synthetic.
XX
PN WO9640772-A2.
XX
PD 19-DEC-1996.
XX
PF 06-JUN-1996; 96WO-US009469.
XX
PR 07-JUN-1995; 95US-00484135.
XX
PA (JOHJ) JOHNSON & JOHNSON.
XX
PI Johnson DL, Zivin RA;
XX
DR WPI; 1997-099920/09.
XX
PT Activating cell surface receptors using peptide dimer agonists - also,
PT new dimers of erythropoietin receptor binding peptide(s) useful for
PT treating patient having disorder characterised by EPO deficiency.
XX
PS Disclosure; Fig 9; 110pp; English.
XX
CC The present peptide is a specific example of a claimed generic monomer
CC subunit of an erythropoietin (EPO) receptor binding dimer, which
CC comprises 2 EPO receptor binding monomers of 10 to 40 amino acids, and
CC activates or improves the bioactivity of the EPO cell surface receptor.
CC The dimer can be used to treat disorders resulting from EPO deficiency by
CC improving the activity of its cell surface receptor, e.g. end stage renal
CC failure/dialysis, anaemia associated with AIDS or chronic inflammatory
CC diseases such as rheumatoid arthritis and chronic bowel inflammation and
CC autoimmune disease. It can also be used to boost the red cell count of a
CC patient prior to surgery or as pretreatment to transfusion. The dimer
CC peptide exhibits increased biological potency in vitro and in vivo
CC relative to its component monomeric agonists. Dimerisation may also
CC convert cell surface receptor antagonists into agonists
XX
SQ Sequence 20 AA;

Query Match 93.4%; Score 57; DB 2; Length 20;
Best Local Similarity 57.1%; Pred. No. 0.022;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 YXCXXGPXTWXCXP 14
|| || || | |
Db 4 YSCRMGPMTWVCSP 17

Search completed: March 31, 2006, 16:22:22
Job time : 55.9801 secs

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: March 31, 2006, 16:22:51 ; Search time 8.70647 Seconds
(without alignments)
154.717 Million cell updates/sec

Title: US-10-609-217-83
Perfect score: 61
Sequence: 1 YXCXXGPXTWXCXP 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	45	73.8	19	1	EWSMAN	ancovenin - Strept
2	40	65.6	318	2	E87929	protein T22H2.6 [i
3	40	65.6	345	2	T25138	hypothetical prote
4	40	65.6	358	2	T25137	hypothetical prote
5	39	63.9	1531	1	DVHUAR	multidrug resistan
6	37	60.7	294	2	S13141	hypothetical prote
7	37	60.7	321	2	F90826	hypothetical prote
8	37	60.7	324	2	G85684	unknown protein en
9	37	60.7	415	2	PC4407	envelope protein -
10	37	60.7	460	2	S06022	regulatory protein
11	37	60.7	475	2	H84137	hypothetical prote
12	37	60.7	652	2	S25265	outer membrane pro
13	37	60.7	652	2	D82317	iron-regulated out
14	37	60.7	3175	1	RRWVEV	genome polyprotein
15	36	59.0	123	2	I52427	guanine-nucleotide
16	36	59.0	123	2	S29714	guanine-nucleotide
17	36	59.0	350	1	DEZPA	alcohol dehydrogen
18	36	59.0	466	2	A36674	transcription fact
19	36	59.0	571	1	S30253	GABA transport pro
20	35.5	58.2	4543	1	A53102	alpha-2-macroglobu
21	35	57.4	341	1	PYVZCB	spheroidin precurs
22	35	57.4	612	2	T36880	hypothetical prote
23	35	57.4	645	2	T27186	hypothetical prote
24	35	57.4	814	2	G02390	disintegrin-like m
25	35	57.4	2531	2	S18188	notch protein homo
26	35	57.4	2531	2	A46019	notch-1 protein -
27	35	57.4	2555	2	A40043	notch protein homo
28	34.5	56.6	1661	2	T31330	head-activator bin
29	34	55.7	19	1	EWSMCN	cinnamycin - Strep
30	34	55.7	78	1	EWSMYG	cinnamycin precurs
31	34	55.7	217	2	E95370	hypothetical prote
32	34	55.7	308	2	S74719	hypothetical prote
33	34	55.7	1472	2	B54774	ATP binding casset
34	34	55.7	1693	2	S76086	beta transducin-li
35	33.5	54.9	4544	1	S02392	alpha-2-macroglobu
36	33.5	54.9	4545	1	S25111	alpha-2-macroglobu
37	33	54.1	68	2	B43940	lactococcin B prec
38	33	54.1	119	2	B98236	exsI protein prote
39	33	54.1	177	2	T01705	hypothetical prote
40	33	54.1	217	2	H86188	protein T25N20.5 [
41	33	54.1	266	2	H86407	F3H9.15 protein -
42	33	54.1	292	2	G88071	protein ZK1240.5 [
43	33	54.1	326	4	S61652	hypothetical prote
44	33	54.1	410	2	S38238	hypothetical prote
45	33	54.1	449	2	AC0234	probable exported

ALIGNMENTS

RESULT 1
EWSMAN
ancovenin - Streptomyces sp. (strain A647P-2)
C;Species: Streptomyces sp.
C;Date: 12-May-1994 #sequence_revision 19-May-1994 #text_change 09-Jul-2004
C;Accession: A61284
R;Wakamiya, T.; Ueki, Y.; Shiba, T.; Kido, Y.; Motoki, Y.
Tetrahedron Lett. 26, 665-668, 1985
A;Title: The structure of ancovenin, a new peptide inhibitor of angiotensin I converting
A;Reference number: A61284
A;Accession: A61284
A;Molecule type: protein
A;Residues: 1-19 <WAK>
A;Cross-references: UNIPROT:P38655; UNIPARC:UPI0000052CC3
C;Superfamily: cinnamycin precursor
C;Keywords: antibiotic; lanthionine
F;1-18/Cross-link: (2S,3S,6R)-3-methyl-lanthionine (Cys-Thr) #status experimental
F;4-14/Cross-link: sn-(2S,6R)-lanthionine (Ser-Cys) #status experimental
F;5-11/Cross-link: (2S,3S,6R)-3-methyl-lanthionine (Cys-Thr) #status experimental
F;6/Modified site: dehydroalanine (Ser) #status experimental

Query Match 73.8%; Score 45; DB 1; Length 19;
Best Local Similarity 60.0%; Pred. No. 0.11;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy          3 CXXGPXTWXC 12
              |  ||  ||  |
Db          5 CSFGPLTWSC 14
```

RESULT 2
E87929
protein T22H2.6 [imported] - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Dec-2002
C;Accession: E87929
R;anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
A;Reference number: A75000; MUID:99069613; PMID:9851916
A;Note: see websites genome.wustl.edu/gsc/C_elegans/ and www_sanger.ac.uk/Projects/C_ele
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A;Accession: E87929
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-318 <STO>
A;Cross-references: UNIPARC:UPI0000177C8F; GB:chr_I; PIDN:CAB04752.1; PID:g3880056; GSPD
C;Genetics:
A;Gene: T22H2.6
A;Map position: 1
C;Superfamily: protein T22H2.6

```
  Query Match             65.6%;  Score 40;  DB 2;  Length 318;
  Best Local Similarity   50.0%;  Pred. No. 9.5;
  Matches    6;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          3 CXXGPXTWXCXP 14
              |  |  || | |
Db         71 CKLGDNTWGCCP 82


RESULT 3
T25138
hypothetical protein T22H2.6b - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T25138
R;Lennard, N.
submitted to the EMBL Data Library, November 1996
A;Reference number: Z19985
A;Accession: T25138
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-345 <WIL>
A;Cross-references: UNIPROT:Q9U362; UNIPARC:UPI000002A1D2; EMBL:Z81595; PIDN:CAB54305.1;
A;Experimental source: clone T22H2
C;Genetics:
A;Gene: CESP:T22H2.6b
A;Map position: 1
A;Introns: 93/3; 232/3; 314/3
C;Superfamily: protein T22H2.6

  Query Match             65.6%;  Score 40;  DB 2;  Length 345;
  Best Local Similarity   50.0%;  Pred. No. 10;
  Matches    6;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          3 CXXGPXTWXCXP 14
              |  |  || | |
Db        111 CKLGDNTWGCCP 122


RESULT 4
T25137
hypothetical protein T22H2.6a - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T25137
R;Lennard, N.
submitted to the EMBL Data Library, November 1996
A;Reference number: Z19985
A;Accession: T25137
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-358 <WIL>
A;Cross-references: UNIPROT:Q9U362; UNIPARC:UPI000008667D; EMBL:Z81595; PIDN:CAB54304.1;
A;Experimental source: clone T22H2
C;Genetics:
A;Gene: CESP:T22H2.6a
A;Map position: 1
A;Introns: 93/3; 232/3; 314/3
C;Superfamily: protein T22H2.6

  Query Match             65.6%;  Score 40;  DB 2;  Length 358;
  Best Local Similarity   50.0%;  Pred. No. 11;
  Matches    6;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          3 CXXGPXTWXCXP 14
              |  |  || | |
Db        111 CKLGDNTWGCCP 122


RESULT 5
DVHUAR
multidrug resistance protein (cell line H69AR) - human
N;Alternate names: multidrug resistance-associated protein (MRP)
C;Species: Homo sapiens (man)
C;Date: 30-Jun-1993 #sequence_revision 05-Dec-1998 #text_change 19-Jan-2001
C;Accession: A44231; A37495
R;Cole, S.P.C.; Bhardwaj, G.; Gerlach, J.H.; Mackie, J.E.; Grant, C.E.; Almquist, K.C.; S
Science 258, 1650-1654, 1992
A;Title: Overexpression of a transporter gene in a multidrug-resistant human lung cancer
A;Reference number: A44231; MUID:93088080; PMID:1360704
A;Accession: A44231
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 'MAPTRSGTGMSRGIPATPTSPSAFRTRSSCGCLVFTSGPV',50-1531 <CO1>
A;Cross-references: UNIPARC:UPI00001746CB; GB:L05628; NID:g1835658
A;Experimental source: small cell lung carcinoma cell line H69AR
A;Note: sequence extracted from NCBI backbone (NCBIP:119851); this sequence has been corr
R;Cole, S.P.C.; Deeley, R.G.
Science 260, 879, 1993
A;Title: Multidrug resistance-associated protein: sequence correction.
A;Reference number: A37495; MUID:93262415; PMID:8098549
A;Accession: A37495
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-60 <CO2>
A;Cross-references: UNIPARC:UPI00001746CC; GB:L05628; NID:g1835658
A;Note: sequence extracted from NCBI backbone (NCBIP:131929)
C;Genetics:
A;Gene: GDB:MRP
A;Cross-references: GDB:136335; OMIM:158343
A;Map position: 16p13.1-16p13.1
C;Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology
C;Keywords: antibiotic resistance; ATP; duplication; nucleotide binding; P-loop; transmem
F;661-844/Domain: ATP-binding cassette homology <ABC1>
F;678-685/Region: nucleotide-binding motif A (P-loop)
F;788-792/Region: nucleotide-binding motif B
F;1310-1503/Domain: ATP-binding cassette homology <ABC2>
F;1327-1334/Region: nucleotide-binding motif A (P-loop)
F;1450-1454/Region: nucleotide-binding motif B

  Query Match             63.9%;  Score 39;  DB 1;  Length 1531;
  Best Local Similarity   42.9%;  Pred. No. 55;
  Matches    6;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          1 YXCXXGPXTWXCXP 14
              |    |  || | |
Db        544 YLSAVGTFTWVCTP 557


RESULT 6
S13141
hypothetical protein (ribosomal RNA repeat region) - Giardia lamblia
C;Species: Giardia lamblia
C;Date: 06-Dec-1996 #sequence_revision 06-Dec-1996 #text_change 05-Oct-2004
C;Accession: S13141; S10886
R;Upcroft, J.A.; Healey, A.; Mitchell, R.; Boreham, P.F.L.; Upcroft, P.
Nucleic Acids Res. 18, 7077-7081, 1990
A;Title: Antigen expression from the ribosomal DNA repeat unit of Giardia intestinalis.
A;Reference number: S13141; MUID:91088287; PMID:2263466
A;Accession: S13141
A;Molecule type: DNA
A;Residues: 1-294 <UPC>
A;Cross-references: UNIPROT:Q9XZV7; UNIPARC:UPI0000177CC5; EMBL:X52949
A;Note: the source is designated as Giardia intestinalis
A;Note: readthrough of the terminator TAG is supposed to occur between residues 241-Ala a
R;Healey, A.; Mitchell, R.; Upcroft, J.A.; Boreham, P.F.L.; Upcroft, P.
Nucleic Acids Res. 18, 4006, 1990
A;Title: Complete nucleotide sequence of the ribosomal RNA tandem repeat unit from Giardi
A;Reference number: S10886; MUID:90326542; PMID:2374731
A;Accession: S10886
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-241 <HEA>
A;Cross-references: UNIPARC:UPI0000177CC6; EMBL:X52949
A;Note: the source is designated as Giardia intestinalis
A;Note: the assignment of the coding region has been revised in reference S13141
```

C;Superfamily: Proline-rich peptide P-B

Query Match 60.7%; Score 37; DB 2; Length 294;
Best Local Similarity 62.5%; Pred. No. 29;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          7 PXTWXCXP 14
              | || | |
Db         93 PRTWACLP 100
```

RESULT 7
F90826
hypothetical protein ECs1582 [imported] - Escherichia coli (strain O157:H7, substrain RI
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: F90826
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: F90826
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-321 <HAY>
A;Cross-references: UNIPROT:Q8X356; UNIPARC:UPI00000D29F6; GB:BA000007; PIDN:BAB35005.1;
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs1582

Query Match 60.7%; Score 37; DB 2; Length 321;
Best Local Similarity 62.5%; Pred. No. 32;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          7 PXTWXCXP 14
              | || | |
Db        179 PETWLCSP 186
```

RESULT 8
G85684
unknown protein encoded by prophage CP-933C [imported] - Escherichia coli (strain O157:H
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: G85684
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: G85684
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-324 <STO>
A;Cross-references: UNIPROT:Q8X3P7; UNIPARC:UPI00000D0ED9; GB:AE005174; NID:g12514761; P
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: Z1842

Query Match 60.7%; Score 37; DB 2; Length 324;
Best Local Similarity 62.5%; Pred. No. 32;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          7 PXTWXCXP 14
              | || | |
Db        179 PETWLCSP 186
```

RESULT 9
PC4407
envelope protein - hepatitis C virus (fragment)
C;Species: hepatitis C virus
C;Date: 10-Nov-1997 #sequence_revision 23-Jan-1998 #text_change 09-Jul-2004
C;Accession: PC4407
R;Li, G.; Yao, J.; Peng, W.
Chinese J. Virol. 13, 24-32, 1997
A;Title: Sequence of genomic region of hepatitis C virus envelope proteins from a Guangd
A;Reference number: PC4407
A;Accession: PC4407
A;Molecule type: genomic RNA
A;Residues: 1-415 <LIA>
A;Cross-references: UNIPROT:Q7LZY4; UNIPARC:UPI0000178545
A;Note: the authors translated the codon ATA for residues 93 and 249 as Met
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: envelope protein

Query Match 60.7%; Score 37; DB 2; Length 415;
Best Local Similarity 41.7%; Pred. No. 40;
Matches 5; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

```
Qy          3 CXXGPXTWXCXP 14
              |   | :| | |
Db        329 CGIVPASWVCGP 340
```

RESULT 10
S06022
regulatory protein O2 - maize
C;Species: Zea mays (maize)
C;Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 31-Dec-2004
C;Accession: S06022; S06009
R;Hartings, H.; Maddaloni, M.; Lazzaroni, N.; di Fonzo, N.; Motto, M.; Salamini, F.; Tho
EMBO J. 8, 2795-2801, 1989
A;Title: The O2 gene which regulates zein deposition in maize endosperm encodes a protei
A;Reference number: S06022; MUID:90059860; PMID:2479535
A;Accession: S06022
A;Molecule type: mRNA
A;Residues: 1-460 <HAR>
A;Cross-references: UNIPROT:P12959; UNIPARC:UPI000016E05D; GB:X16618; NID:g22383; PIDN:C
R;Maddaloni, M.; di Fonzo, N.; Hartings, H.; Lazzaroni, N.; Salamini, F.; Thompson, R.;
Nucleic Acids Res. 17, 7532, 1989
A;Title: The sequence of the zein regulatory gene opaque-2 (02) of Zea Mays.
A;Reference number: S06009; MUID:90016825; PMID:2798113
A;Accession: S06009
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-22,29-149,'D',151-460 <MAD>
A;Cross-references: UNIPARC:UPI00001794F4; EMBL:X15544
C;Genetics:
A;Gene: opaque 2
A;Map position: 7
A;Introns: 148/3; 168/3; 238/2; 263/3; 305/3
C;Superfamily: BZIP protein; fos/jun DNA-binding domain homology
C;Keywords: DNA binding; nucleus; transcription regulation
F;227-267/Domain: fos/jun DNA-binding domain homology <FJD>

Query Match 60.7%; Score 37; DB 2; Length 460;
Best Local Similarity 71.4%; Pred. No. 43;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          6 GPXTWXC 12
              || || |
Db        436 GPYTWTC 442
```

RESULT 11
H84137
hypothetical protein BH3904 [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C;Accession: H84137
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132

```
A;Accession: H84137
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-475 <STO>
A;Cross-references: UNIPROT:Q9K628; UNIPARC:UPI00000C432F; GB:AP001520; GB:BA000004; NID
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH3904

  Query Match             60.7%;  Score 37;  DB 2;  Length 475;
  Best Local Similarity   62.5%;  Pred. No. 45;
  Matches    5;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          3 CXXGPXTW 10
              |  || ||
Db        156 CAGGPSTW 163


RESULT 12
S25265
outer membrane protein irgA precursor - Vibrio cholerae
N;Alternate names: ferrienterochelin receptor homolog
C;Species: Vibrio cholerae
C;Date: 28-May-1993 #sequence_revision 28-May-1993 #text_change 09-Jul-2004
C;Accession: S25265; A37834
R;Goldberg, M.B.; Boyko, S.A.; Butterton, J.R.; Stoebner, J.A.; Payne, S.M.; Calderwood,
Mol. Microbiol. 6, 2407-2418, 1992
A;Title: Characterization of a Vibrio cholerae virulence factor homologous to the family
A;Reference number: S25265; MUID:93023868; PMID:1406279
A;Accession: S25265
A;Molecule type: DNA
A;Residues: 1-652 <GOL>
A;Cross-references: UNIPROT:P27772; UNIPARC:UPI0000148DB5; GB:U72152; EMBL:M63192; NID:g
A;Note: the sequence from Fig. 3 is inconsistent with that from Fig. 2 in having 299-Thr
R;Goldberg, M.B.; Boyko, S.A.; Calderwood, S.B.
J. Bacteriol. 172, 6863-6870, 1990
A;Title: Transcriptional regulation by iron of a Vibrio cholerae virulence gene and homo
A;Reference number: A37834; MUID:91072235; PMID:2174861
A;Accession: A37834
A;Molecule type: DNA
A;Residues: 1-152,'D' <GO2>
A;Cross-references: UNIPARC:UPI000017838A; GB:M37773
C;Genetics:
A;Gene: irgA
C;Superfamily: ferrienterochelin receptor; tonB-dependent receptor amino-terminal homolo
C;Keywords: membrane protein
F;1-25/Domain: signal sequence #status predicted <SIG>
F;26-652/Product: outer membrane protein irgA #status predicted <MAT>
F;68-214/Domain: tonB-dependent receptor amino-terminal homology <TNN>
F;367-652/Domain: tonB-dependent receptor carboxyl-terminal homology <TNC>

  Query Match             60.7%;  Score 37;  DB 2;  Length 652;
  Best Local Similarity   41.7%;  Pred. No. 59;
  Matches    5;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          3 CXXGPXTWXCXP 14
              |  || |   |
Db        492 CTAGPNQWGATP 503


RESULT 13
D82317
iron-regulated outer membrane virulence protein, TonB receptor family VC0475 [imported]
C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: D82317
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A;Reference number: A82035; MUID:20406833; PMID:10952301
A;Accession: D82317
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-652 <HEI>
A;Cross-references: UNIPROT:P27772; UNIPARC:UPI000012D88F; GB:AE004134; GB:AE003852; NID:
A;Experimental source: serogroup O1; strain N16961; biotype El Tor
C;Genetics:
A;Gene: VC0475
A;Map position: 1
C;Superfamily: ferrienterochelin receptor; tonB-dependent receptor amino-terminal homolog

  Query Match             60.7%;  Score 37;  DB 2;  Length 652;
  Best Local Similarity   41.7%;  Pred. No. 59;
  Matches    5;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          3 CXXGPXTWXCXP 14
              |  || |   |
Db        492 CTAGPNQWGAAP 503


RESULT 14
RRWVEV
genome polyprotein - equine arteritis virus
N;Contains: RNA-directed RNA polymerase (EC 2.7.7.48)
C;Species: equine arteritis virus
A;Note: host Equus caballus (domestic horse)
C;Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 09-Jul-2004
C;Accession: A39925; S10158; B39925
R;Den Boon, J.A.; Snijder, E.J.; Chirnside, E.D.; De Vries, A.A.F.; Horzinek, M.C.; Spaar
J. Virol. 65, 2910-2920, 1991
A;Title: Equine arteritis virus is not a togavirus but belongs to the coronaviruslike sup
A;Reference number: A39925; MUID:91237805; PMID:1851863
A;Accession: A39925
A;Molecule type: genomic RNA
A;Residues: 1-3175 <DEN>
A;Cross-references: UNIPROT:P19811; UNIPARC:UPI0000134685; EMBL:X53459
A;Note: a -1 ribosomal frameshift occurs between the codons AAC for 1727-Asn and CUG for
R;de Vries, A.A.F.; Chirnside, E.D.; Bredenbeek, P.J.; Gravestein, L.A.; Horzinek, M.C.;
Nucleic Acids Res. 18, 3241-3247, 1990
A;Title: All subgenomic mRNAs of equine arteritis virus contain a common leader sequence.
A;Reference number: S10158; MUID:90287699; PMID:2162519
A;Accession: S10158
A;Status: translation not shown
A;Molecule type: genomic RNA
A;Residues: 1-17 <VRI>
A;Cross-references: UNIPARC:UPI0000172725; EMBL:X52277
C;Superfamily: equine arteritis virus RNA-directed RNA polymerase
C;Keywords: nucleotidyltransferase

  Query Match             60.7%;  Score 37;  DB 1;  Length 3175;
  Best Local Similarity   35.7%;  Pred. No. 2.3e+02;
  Matches    5;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          1 YXCXXGPXTWXCXP 14
              | |      | | |
Db        242 YVCDISEADWSCLP 255


RESULT 15
I52427
guanine-nucleotide-releasing protein Mss4 - human
C;Species: Homo sapiens (man)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C;Accession: I52427
R;Yu, H.; Schreiber, S.L.
Biochemistry 34, 9103-9110, 1995
A;Title: Cloning, Zn2+ binding, and structural characterization of the guanine nucleotide
A;Reference number: I52427; MUID:95345082; PMID:7619808
A;Accession: I52427
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-123 <RES>
A;Cross-references: UNIPROT:P47224; UNIPARC:UPI00001117CC; GB:S78873; NID:g1037135; PIDN
C;Genetics:
```

A;Gene: GDB:MSS4
A;Cross-references: GDB:683578

Query Match 59.0%; Score 36; DB 2; Length 123;
Best Local Similarity 50.0%; Pred. No. 21;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy          3 CXXGPXTWXC 12
              |  ||  | |
Db         97 CEIGPIGWHC 106
```

Search completed: March 31, 2006, 16:37:12
Job time : 10.7065 secs

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: March 31, 2006, 16:09:36 ; Search time 52.4478 Seconds
(without alignments)
188.328 Million cell updates/sec

Title: US-10-609-217-83
Perfect score: 61
Sequence: 1 YXCXXGPXTWXCXP 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	46	75.4	352	2	Q4IMN3_GIBZE	Q4imn3 gibberella
2	45	73.8	19	1	DURC_STRGP	P36503 streptomyce
3	45	73.8	19	1	LANC_STRS6	P38655 streptomyce
4	43	70.5	378	1	TAZ_DROME	Q9v6g5 drosophila
5	43	70.5	414	2	Q4SAV9_TETNG	Q4sav9 tetraodon n
6	41	67.2	532	2	Q8WWV6_HUMAN	Q8wwv6 homo sapien
7	41	67.2	534	2	Q96SA2_HUMAN	Q96sa2 homo sapien
8	41	67.2	577	2	Q5REH9_PONPY	Q5reh9 pongo pygma
9	41	67.2	589	2	Q5R770_PONPY	Q5r770 pongo pygma
10	40.5	66.4	499	2	Q6ARY7_DESPS	Q6ary7 desulfotale
11	40	65.6	167	2	Q6ZW93_HUMAN	Q6zw93 homo sapien
12	40	65.6	173	2	Q5VHX3_EAV	Q5vhx3 equine arte
13	40	65.6	180	2	Q4I355_GIBZE	Q4i355 gibberella
14	40	65.6	345	2	Q7JKP2_CAEEL	Q7jkp2 caenorhabdi
15	40	65.6	358	2	Q9U362_CAEEL	Q9u362 caenorhabdi
16	40	65.6	698	2	Q810G8_RAT	Q810g8 rattus norv
17	40	65.6	1523	2	Q810G9_RAT	Q810g9 rattus norv
18	40	65.6	1525	2	Q5F364_CHICK	Q5f364 gallus gall
19	40	65.6	1528	1	MRP1_MOUSE	O35379 mus musculu
20	40	65.6	1530	2	Q8HXQ5_BOVIN	Q8hxq5 bos taurus
21	40	65.6	1531	2	Q6UR05_CANFA	Q6ur05 canis famil
22	40	65.6	1531	2	Q864R9_MACFA	Q864r9 macaca fasc
23	40	65.6	1531	2	Q864S0_MACFA	Q864s0 macaca fasc
24	40	65.6	1532	2	Q810E4_RAT	Q810e4 rattus norv
25	40	65.6	1532	2	Q8CG09_RAT	Q8cg09 rattus norv
26	40	65.6	2022	2	Q61J27_CAEBR	Q61j27 caenorhabdi
27	39	63.9	172	2	Q6ZT75_HUMAN	Q6zt75 homo sapien
28	39	63.9	172	2	Q6ZWC2_HUMAN	Q6zwc2 homo sapien
29	39	63.9	173	2	Q9WD22_EAV	Q9wd22 equine arte
30	39	63.9	373	2	O70709_9POXV	O70709 anomala cup
31	39	63.9	691	2	O43333_HUMAN	O43333 homo sapien
32	39	63.9	1215	2	Q68CP7_HUMAN	Q68cp7 homo sapien
33	39	63.9	1400	2	Q9UQ98_HUMAN	Q9uq98 homo sapien
34	39	63.9	1439	2	Q59GI9_HUMAN	Q59gi9 homo sapien
35	39	63.9	1456	2	Q9UQA0_HUMAN	Q9uqa0 homo sapien
36	39	63.9	1459	2	Q9UQ97_HUMAN	Q9uq97 homo sapien
37	39	63.9	1515	2	Q9UQ99_HUMAN	Q9uq99 homo sapien
38	39	63.9	1531	1	MRP1_HUMAN	P33527 homo sapien
39	38	62.3	61	2	O70227_RAT	O70227 rattus norv
40	38	62.3	167	2	Q651J0_ORYSA	Q651j0 oryza sativ
41	38	62.3	285	2	Q8NAV2_HUMAN	Q8nav2 homo sapien
42	38	62.3	329	2	Q7Z758_HUMAN	Q7z758 homo sapien
43	38	62.3	336	2	Q5PR99_BRARE	Q5pr99 brachydanio
44	38	62.3	967	2	Q59FS0_HUMAN	Q59fs0 homo sapien
45	37.5	61.5	2465	2	Q4RXZ7_TETNG	Q4rxz7 tetraodon n

ALIGNMENTS

RESULT 1
Q4IMN3_GIBZE
ID Q4IMN3_GIBZE PRELIMINARY; PRT; 352 AA.
AC Q4IMN3;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=FG01525.1;
OS Gibberella zeae PH-1.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
OX NCBI_TaxID=229533;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PH-1;
RA Birren B., Nusbaum C., Abouelleil A., Allen N., Anderson S.,
RA Arachchi H.M., Barna N., Bastien V., Bloom T., Boguslavkiy L.,
RA Boukhgalter B., Butler J., Calvo S.E., Camarata J., Chang J.,
RA Choepel Y., Collymore A., Cook A., Cooke P., Corum B., DeArellano K.,
RA Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
RA Erickson J., Faro S., Ferreira P., FitzGerald M., Gage D., Galagan J.,
RA Gardyna S., Gnerre S., Graham L., Grand-Pierre N., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Horton L., Hulme W., Iliev I.,
RA Jaffe D., Johnson R., Jones C., Kamal M., Kamat A., Karatas A.,
RA Kells C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
RA Ma L.-J., Mabbitt R., MacLean C., Macdonald P., Major J., Manning J.,
RA Matthews C., Mauceli E., McCarthy M., Meldrim J., Meneus L.,
RA Mihova T., Mlenga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
RA Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neil D.,
RA Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,
RA Rachupka A., Ramasamy U., Raymond C., Retta R., Rise C., Rogov P.,
RA Roman J., Schauer S., Schupback R., Seaman S., Severy P., Smirnov S.,
RA Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,
RA Talamas J., Tesfaye S., Theodore J., Topham K., Travers M.,
RA Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,
RA Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
RA Lander E.;
RT "Fusarium graminearum genome sequence.";
RL Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AACM01000077; EAA68151.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 352 AA; 38308 MW; 670BA49FC645A7F8 CRC64;

Query Match 75.4%; Score 46; DB 2; Length 352;
Best Local Similarity 50.0%; Pred. No. 4.4;
Matches 6; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
Qy          3 CXXGPXTWXCXP 14
              |  | || | |
Db        184 CTSNPSTWRCYP 195
```

```
RESULT 2
DURC_STRGP
ID   DURC_STRGP     STANDARD;      PRT;    19 AA.
AC   P36503;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Lantibiotic duramycin C.
OS   Streptomyces griseoluteus.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=29306;
RN   [1]
RP   PROTEIN SEQUENCE.
RC   STRAIN=R2107;
RX   MEDLINE=91107436; PubMed=2125590;
RA   Fredenhagen A., Fendrich G., Marki F., Marki W., Gruner J.,
RA   Raschdorf F., Peter H.H.;
RT   "Duramycins B and C, two new lanthionine containing antibiotics as
RT   inhibitors of phospholipase A2. Structural revision of duramycin and
RT   cinnamycin.";
RL   J. Antibiot. 43:1403-1412(1990).
RN   [2]
RP   STRUCTURE BY NMR.
RA   Zimmermann N., Freund S., Fredenhagen A., Jung G.;
RT   "Solution structure of the lantibiotics duramycin B and C.";
RL   (In) Schneider C.H., Eberles A.N. (eds.);
RL   Peptides 1992, pp.519-520, Escom Science Publishers, Leiden (1993).
RN   [3]
RP   STRUCTURE BY NMR.
RX   MEDLINE=93387292; PubMed=8375380;
RA   Zimmermann N., Freund S., Fredenhagen A., Jung G.;
RT   "Solution structures of the lantibiotics duramycin B and C.";
RL   Eur. J. Biochem. 216:419-428(1993).
CC   -!- FUNCTION: Acts as inhibitor of phospholipase A2.
CC   -!- PTM: Maturation of lantibiotics involves the enzymic conversion of
CC       Thr, and Ser into dehydrated AA and the formation of thioether
CC       bonds with cysteine or the formation of dialkylamine bonds with
CC       lysine. This is followed by membrane translocation and cleavage of
CC       the modified precursor.
CC   -!- SIMILARITY: Belongs to the type B lantibiotic family.
CC   ----------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ----------------------------------------------------------------------
KW   Antibiotic; Antimicrobial; Bacteriocin; Direct protein sequencing;
KW   Lantibiotic; Thioether bond.
FT   CROSSLNK      1     18       Beta-methyllanthionine (Cys-Thr).
FT   CROSSLNK      4     14       Lanthionine (Ser-Cys).
FT   CROSSLNK      5     11       Beta-methyllanthionine (Cys-Thr).
FT   CROSSLNK      6     19       Lysinoalanine (Ser-Lys).
SQ   SEQUENCE   19 AA;  2007 MW;  E2404ECE3F95286A CRC64;

  Query Match             73.8%;  Score 45;  DB 1;  Length 19;
  Best Local Similarity   60.0%;  Pred. No. 0.44;
  Matches    6;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          3 CXXGPXTWXC 12
              |  ||  || |
Db          5 CSYGPLTWSC 14


RESULT 3
LANC_STRS6
ID   LANC_STRS6     STANDARD;      PRT;    19 AA.
AC   P38655;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Lantibiotic ancovenin.
OS   Streptomyces sp. (strain A647P-2).
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=72591;
RN   [1]
RP   PROTEIN SEQUENCE.
RA   Wakamiya T., Ueki Y., Shiba T., Kido Y., Motoki Y.;
RT   "The structure of ancovenin, a new peptide inhibitor of angiotensin I
RT   converting enzyme.";
RL   Tetrahedron Lett. 26:665-668(1985).
CC   -!- FUNCTION: Acts as an inhibitor of angiotensin I converting enzyme.
CC   -!- PTM: Maturation of lantibiotics involves the enzymic conversion of
CC       Thr, and Ser into dehydrated AA and the formation of thioether
CC       bonds with cysteine or the formation of dialkylamine bonds with
CC       lysine. This is followed by membrane translocation and cleavage of
CC       the modified precursor.
CC   -!- SIMILARITY: Belongs to the type B lantibiotic family.
CC   ----------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ----------------------------------------------------------------------
DR   PIR; A61284; EWSMAN.
KW   Antibiotic; Antimicrobial; Bacteriocin; Direct protein sequencing;
KW   Lantibiotic; Thioether bond.
FT   CROSSLNK      1     18       Beta-methyllanthionine (Cys-Thr).
FT   CROSSLNK      4     14       Lanthionine (Ser-Cys).
FT   CROSSLNK      5     11       Beta-methyllanthionine (Cys-Thr).
FT   CROSSLNK      6     19       Lysinoalanine (Ser-Lys).
SQ   SEQUENCE   19 AA;  2033 MW;  F434299E2736286A CRC64;

  Query Match             73.8%;  Score 45;  DB 1;  Length 19;
  Best Local Similarity   60.0%;  Pred. No. 0.44;
  Matches    6;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          3 CXXGPXTWXC 12
              |  ||  || |
Db          5 CSFGPLTWSC 14


RESULT 4
TAZ_DROME
ID   TAZ_DROME      STANDARD;      PRT;   378 AA.
AC   Q9V6G5; Q8ML32; Q8SZ79; Q9U9U8; Q9V6G4;
DT   28-FEB-2003 (Rel. 41, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Tafazzin homolog.
GN   Name=tafazzin; ORFNames=CG8766;
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   NUCLEOTIDE SEQUENCE (ISOFORM A).
RC   STRAIN=Canton-S;
RA   Benevolenskaya E.V., Frolov M.V., Birchler J.A.;
RT   "Drosophila homolog of the human G4.5 gene encoding tafazzin
RT   proteins.";
RL   Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC   STRAIN=Berkeley;
RX   MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
```

```
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [3]
RP   GENOME REANNOTATION, AND ALTERNATIVE SPLICING.
RX   MEDLINE=22426069; PubMed=12537572;
RA   Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA   Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA   Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA   Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA   Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA   Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA   Lewis S.E.;
RT   "Annotation of the Drosophila melanogaster euchromatic genome: a
RT   systematic review.";
RL   Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN   [4]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM A).
RC   STRAIN=Berkeley; TISSUE=Embryo;
RX   MEDLINE=22426066; PubMed=12537569;
RA   Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,
RA   George R.A., Guarin H., Kronmiller B., Pacleb J.M., Park S., Wan K.H.,
RA   Rubin G.M., Celniker S.E.;
RT   "A Drosophila full-length cDNA resource.";
RL   Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).
RN   [5]
RP   NUCLEOTIDE SEQUENCE OF 1-39 (ISOFORM A).
RC   STRAIN=Canton-S;
RX   MEDLINE=20556141; PubMed=11102369;
RA   Frolov M.V., Benevolenskaya E.V., Birchler J.A.;
RT   "The oxen gene of Drosophila encodes a homolog of subunit 9 of yeast
RT   ubiquinol-cytochrome c oxidoreductase complex: evidence for modulation
RT   of gene expression in response to mitochondrial activity.";
RL   Genetics 156:1727-1736(2000).
CC   -!- INTERACTION:
CC       Q9VA96:CG15530; NbExp=1; IntAct=EBI-244104, EBI-151361;
CC   -!- SUBCELLULAR LOCATION: Isoforms with hydrophobic N-terminus are
CC       thought to be membrane-anchored (By similarity).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=3;
CC       Name=A;
CC         IsoId=Q9V6G5-1; Sequence=Displayed;
CC       Name=B;
CC         IsoId=Q9V6G5-2; Sequence=VSP_004450;
CC         Note=No experimental confirmation available;
CC       Name=C;
CC         IsoId=Q9V6G5-3; Sequence=VSP_007017;
CC         Note=No experimental confirmation available;
CC   -!- DOMAIN: The hydrophilic domain may serve as an exposed loop
CC       interacting with other proteins (By similarity).
CC   -!- SIMILARITY: Belongs to the taffazin family.
CC   -!- CAUTION: Ref.1 sequence differs from that shown due to a
CC       frameshift in position 117.
CC   ---------------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------------
DR   EMBL; AF148684; AAD48409.1; ALT_FRAME; mRNA.
DR   EMBL; AE003821; AAF58461.2; -; Genomic_DNA.
DR   EMBL; AE003821; AAF58462.3; -; Genomic_DNA.
DR   EMBL; AE003821; AAM68652.2; -; Genomic_DNA.
DR   EMBL; AY071059; AAL48681.1; -; mRNA.
DR   EMBL; AF017783; -; NOT_ANNOTATED_CDS; Genomic_DNA.
DR   IntAct; Q9V6G5; -.
DR   Ensembl; CG8766; Drosophila melanogaster.
DR   FlyBase; FBgn0026619; tafazzin.
DR   GO; GO:0016021; C:integral to membrane; NAS.
DR   InterPro; IPR002123; Acyltransferase.
DR   InterPro; IPR000872; Tafazzin.
DR   PANTHER; PTHR12497; Tafazzin; 1.
DR   Pfam; PF01553; Acyltransferase; 1.
DR   PRINTS; PR00979; TAFAZZIN.
KW   Alternative splicing; Transmembrane.
FT   TRANSMEM    138    158       Potential.
FT   REGION      243    283       Hydrophilic (By similarity).
FT   VARSPLIC      1    100       Missing (in isoform B).
FT                                /FTId=VSP_004450.
FT   VARSPLIC      1     41       MFMVVCSNLRRPGHVGAASAARNINWLISEGYTPPIRAMAR
FT                                -> M (in isoform C).
FT                                /FTId=VSP_007017.
SQ   SEQUENCE   378 AA;  43016 MW;  88690B6E2731FFB9 CRC64;

  Query Match             70.5%;  Score 43;  DB 1;  Length 378;
  Best Local Similarity   42.9%;  Pred. No. 16;
  Matches    6;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          1 YXCXXGPXTWXCXP 14
              | |   |  | | |
Db        189 YSCFDDPGLWGCLP 202


RESULT 5
Q4SAV9_TETNG
ID   Q4SAV9_TETNG PRELIMINARY;      PRT;   414 AA.
AC   Q4SAV9;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Chromosome 3 SCAF14679, whole genome shotgun sequence.
GN   ORFNames=GSTENG00021242001;
OS   Tetraodon nigroviridis (Green puffer).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC   Tetradontoidea; Tetraodontidae; Tetraodon.
OX   NCBI_TaxID=99883;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA   Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA   Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA   Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
```

```
RA   Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA   Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA   Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA   Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA   Kellis M., Volff JN., Guigo R., Zody M.C., Mesirov J.,
RA   Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA   Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA   Wincker P., Lander E.S., Weissenbach J., Roest Crollius H.;
RT   "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT   the early vertebrate proto-karyotype.";
RL   Nature 431:946-957(2004).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RG   Genoscope; Whitehead Institute Centre for Genome Research;
RL   Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; CAAE01014679; CAG02223.1; -; Genomic_DNA.
SQ   SEQUENCE   414 AA;  45368 MW;  0522D03EA381377E CRC64;

  Query Match            70.5%;  Score 43;  DB 2;  Length 414;
  Best Local Similarity  50.0%;  Pred. No. 17;
  Matches    6;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          3 CXXGPXTWXCXP 14
              |   | || | |
Db        155 CRMSPSTWGCCP 166


RESULT 6
Q8WWV6_HUMAN
ID   Q8WWV6_HUMAN PRELIMINARY;      PRT;   532 AA.
AC   Q8WWV6;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Fc alpha/mu receptor.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=21638011; PubMed=11779189;
RA   McDonald K.J., Cameron A.J.M., Allen J.M., Jardine A.G.;
RT   "Expression of Fc alpha/mu receptor by human mesangial cells: a
RT   candidate receptor for immune complex deposition in IgA nephropathy.";
RL   Biochem. Biophys. Res. Commun. 290:438-442(2002).
DR   EMBL; AY063125; AAL51154.1; -; mRNA.
DR   Ensembl; ENSG00000162897; Homo sapiens.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR007110; Ig-like.
DR   SMART; SM00409; IG; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
KW   Immunoglobulin domain; Receptor.
SQ   SEQUENCE   532 AA;  57144 MW;  D347A23C0F41EED3 CRC64;

  Query Match            67.2%;  Score 41;  DB 2;  Length 532;
  Best Local Similarity  50.0%;  Pred. No. 50;
  Matches    6;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 YXCXXGPXTWXC 12
              | |  ||  | |
Db         96 YWCRLGPPRWIC 107


RESULT 7
Q96SA2_HUMAN
ID   Q96SA2_HUMAN PRELIMINARY;      PRT;   534 AA.
AC   Q96SA2;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   FKSG87 protein.
GN   Name=FKSG87;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Wang Y.-G., Gong L.;
RL   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF354295; AAK39522.1; -; mRNA.
DR   Ensembl; ENSG00000162897; Homo sapiens.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR007110; Ig-like.
DR   SMART; SM00409; IG; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
KW   Immunoglobulin domain.
SQ   SEQUENCE   534 AA;  56749 MW;  6EF8050E412AF91C CRC64;

  Query Match            67.2%;  Score 41;  DB 2;  Length 534;
  Best Local Similarity  50.0%;  Pred. No. 50;
  Matches    6;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 YXCXXGPXTWXC 12
              | |  ||  | |
Db        116 YWCRLGPPRWIC 127


RESULT 8
Q5REH9_PONPY
ID   Q5REH9_PONPY PRELIMINARY;      PRT;   577 AA.
AC   Q5REH9;
DT   01-FEB-2005 (TrEMBLrel. 29, Created)
DT   01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT   01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE   Hypothetical protein DKFZp469K1129.
GN   Name=DKFZp469K1129;
OS   Pongo pygmaeus (Orangutan).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Pongo.
OX   NCBI_TaxID=9600;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Kidney;
RG   The German cDNA Consortium;
RA   Ottenwaelder B., Obermaier B., Deutschenbaur S., Schaipp A.,
RA   Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL   Submitted (NOV-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; CR857549; CAH89828.1; -; mRNA.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR007110; Ig-like.
DR   SMART; SM00409; IG; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
KW   Hypothetical protein; Immunoglobulin domain.
SQ   SEQUENCE   577 AA;  62062 MW;  AA0FCEE7AB9C4BCD CRC64;

  Query Match            67.2%;  Score 41;  DB 2;  Length 577;
  Best Local Similarity  50.0%;  Pred. No. 54;
  Matches    6;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 YXCXXGPXTWXC 12
              | |  ||  | |
Db        129 YWCRLGPPRWIC 140


RESULT 9
Q5R770_PONPY
ID   Q5R770_PONPY PRELIMINARY;      PRT;   589 AA.
```

```
AC   Q5R770;
DT   01-FEB-2005 (TrEMBLrel. 29, Created)
DT   01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT   01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE   Hypothetical protein DKFZp469A0319.
GN   Name=DKFZp469A0319;
OS   Pongo pygmaeus (Orangutan).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Pongo.
OX   NCBI_TaxID=9600;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Kidney;
RG   The German cDNA Consortium;
RA   Poustka A., Albert R., Moosmayer P., Schupp I., Wellenreuther R.,
RA   Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL   Submitted (NOV-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; CR860248; CAH92390.1; -; mRNA.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR007110; Ig-like.
DR   SMART; SM00409; IG; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
KW   Hypothetical protein; Immunoglobulin domain.
SQ   SEQUENCE   589 AA;  63435 MW;  255BF0FEAACCA812 CRC64;

  Query Match             67.2%;  Score 41;  DB 2;  Length 589;
  Best Local Similarity   50.0%;  Pred. No. 55;
  Matches    6;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 YXCXXGPXTWXC 12
              | |  ||  | |
Db        141 YWCRLGPPRWIC 152


RESULT 10
Q6ARY7_DESPS
ID   Q6ARY7_DESPS PRELIMINARY;      PRT;   499 AA.
AC   Q6ARY7;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Related to cytochrome-c3 hydrogenase (NiFeSe), large subunit.
GN   OrderedLocusNames=DP0159;
OS   Desulfotalea psychrophila.
OC   Bacteria; Proteobacteria; Deltaproteobacteria; Desulfobacterales;
OC   Desulfobulbaceae; Desulfotalea.
OX   NCBI_TaxID=84980;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=LSv54 / DSM 12343;
RX   PubMed=15305914; DOI=10.1111/j.1462-2920.2004.00665.x;
RA   Rabus R., Ruepp A., Frickey T., Rattei T., Fartmann B., Stark M.,
RA   Bauer M., Zibat A., Lombardot T., Becker I., Amann J., Gellner K.,
RA   Teeling H., Leuschner W.D., Gloeckner F.-O., Lupas A.N., Amann R.,
RA   Klenk H.-P.;
RT   "The genome of Desulfotalea psychrophila, a sulfate-reducing bacterium
RT   from permanently cold Arctic sediments.";
RL   Environ. Microbiol. 6:887-902(2004).
DR   EMBL; CR522870; CAG34888.1; -; Genomic_DNA.
DR   GO; GO:0008901; F:ferredoxin hydrogenase activity; IEA.
DR   GO; GO:0046872; F:metal ion binding; IEA.
DR   GO; GO:0016151; F:nickel ion binding; IEA.
DR   GO; GO:0016491; F:oxidoreductase activity; IEA.
DR   GO; GO:0006118; P:electron transport; IEA.
DR   InterPro; IPR001501; Ni_hdL.
DR   Pfam; PF00374; NiFeSe_Hases; 1.
DR   PROSITE; PS00507; NI_HGENASE_L_1; 1.
DR   PROSITE; PS00508; NI_HGENASE_L_2; 1.
KW   Complete proteome; Metal-binding; Nickel; Oxidoreductase.
SQ   SEQUENCE   499 AA;  55328 MW;  8DC870ABF65B7618 CRC64;

  Query Match             66.4%;  Score 40.5;  DB 2;  Length 499;
  Best Local Similarity   50.0%;  Pred. No. 58;
  Matches    7;  Conservative    0;  Mismatches    6;  Indels    1;  Gaps    1;

Qy          1 YXCXXGPXTWXCXP 14
              | |  ||  | || | |
Db        443 YECIV-PTTWNCSP 455


RESULT 11
Q6ZW93_HUMAN
ID   Q6ZW93_HUMAN PRELIMINARY;      PRT;   167 AA.
AC   Q6ZW93;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Hypothetical protein FLJ41423.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Hippocampus;
RA   Kawakami B., Sugiyama A., Takemoto M., Sugiyama T., Irie R.,
RA   Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA   Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA   Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,
RA   Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A., Suzuki Y.,
RA   Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
RL   Submitted (JUL-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK123417; BAC85611.1; -; mRNA.
SQ   SEQUENCE   167 AA;  17960 MW;  266132D59393C276 CRC64;

  Query Match             65.6%;  Score 40;  DB 2;  Length 167;
  Best Local Similarity   50.0%;  Pred. No. 26;
  Matches    5;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          3 CXXGPXTWXC 12
              |  ||  | |
Db         83 CRQGPSVWSC 92


RESULT 12
Q5VHX3_EAV
ID   Q5VHX3_EAV  PRELIMINARY;      PRT;   173 AA.
AC   Q5VHX3;
DT   01-FEB-2005 (TrEMBLrel. 29, Created)
DT   01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT   01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE   Large envelope protein (Fragment).
GN   Name=ORF5;
OS   Equine arteritis virus (EAV).
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC   Arteriviridae; Arterivirus.
OX   NCBI_TaxID=11047;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=S4;
RA   Mittelholzer C., Johansson I., Baule C., Hannant D., Paton D.,
RA   Autorino G.L., Nowotny N., Belak S.;
RT   "Extended phylogeny of equine arteritis virus: division into new
RT   subgroups.";
RL   Submitted (OCT-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY453342; AAS17004.1; -; Genomic_RNA.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   InterPro; IPR001332; Arteri_glycop.
DR   InterPro; IPR003241; EAV_ORF5.
DR   Pfam; PF00951; Arteri_Gl; 1.
DR   ProDom; PD002371; EAV_ORF5; 1.
KW   Envelope protein.
FT   NON_TER       1      1
FT   NON_TER     173    173
```

```
SQ   SEQUENCE   173 AA;  19488 MW;  9147CBD1DD750ADE CRC64;

  Query Match             65.6%;  Score 40;  DB 2;  Length 173;
  Best Local Similarity   41.7%;  Pred. No. 27;
  Matches    5;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 YXCXXGPXTWXC 12
              | |   |  | ||
Db          5 YNCSASPTCWYC 16


RESULT 13
Q4I355_GIBZE
ID   Q4I355_GIBZE PRELIMINARY;      PRT;   180 AA.
AC   Q4I355;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Predicted protein.
GN   ORFNames=FG08353.1;
OS   Gibberella zeae PH-1.
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC   Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
OX   NCBI_TaxID=229533;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=PH-1;
RA   Birren B., Nusbaum C., Abouelleil A., Allen N., Anderson S.,
RA   Arachchi H.M., Barna N., Bastien V., Bloom T., Boguslavkiy L.,
RA   Boukhgalter B., Butler J., Calvo S.E., Camarata J., Chang J.,
RA   Choepel Y., Collymore A., Cook A., Cooke P., Corum B., DeArellano K.,
RA   Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
RA   Erickson J., Faro S., Ferreira P., FitzGerald M., Gage D., Galagan J.,
RA   Gardyna S., Gnerre S., Graham L., Grand-Pierre N., Hafez N.,
RA   Hagopian D., Hagos B., Hall J., Horton L., Hulme W., Iliev I.,
RA   Jaffe D., Johnson R., Jones C., Kamal M., Kamat A., Karatas A.,
RA   Kells C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
RA   Ma L.-J., Mabbitt R., MacLean C., Macdonald P., Major J., Manning J.,
RA   Matthews C., Mauceli E., McCarthy M., Meldrim J., Meneus L.,
RA   Mihova T., Mlenga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
RA   Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neil D.,
RA   Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,
RA   Rachupka A., Ramasamy U., Raymond C., Retta R., Rise C., Rogov P.,
RA   Roman J., Schauer S., Schupback R., Seaman S., Severy P., Smirnov S.,
RA   Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,
RA   Talamas J., Tesfaye S., Theodore J., Topham K., Travers M.,
RA   Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,
RA   Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
RA   Lander E.;
RT   "Fusarium graminearum genome sequence.";
RL   Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; AACM01000335; EAA72141.1; -; Genomic_DNA.
SQ   SEQUENCE   180 AA;  20463 MW;  94C7B5242FEE6ED9 CRC64;

  Query Match             65.6%;  Score 40;  DB 2;  Length 180;
  Best Local Similarity   41.7%;  Pred. No. 28;
  Matches    5;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 YXCXXGPXTWXC 12
              : |  ||  | ||
Db         99 HNCSFGPAPWEC 110


RESULT 14
Q7JKP2_CAEEL
ID   Q7JKP2_CAEEL PRELIMINARY;      PRT;   345 AA.
AC   Q7JKP2;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Hypothetical protein T22H2.6b.
GN   ORFNames=T22H2.6, T22H2.6B;
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC   STRAIN=Bristol N2;
RX   MEDLINE=99069613; PubMed=9851916;
RG   The C. elegans sequencing consortium;
RT   "Genome sequence of the nematode C. elegans: a platform for
RT   investigating biology.";
RL   Science 282:2012-2018(1998).
DR   EMBL; Z81595; CAB54305.1; -; Genomic_DNA.
DR   Ensembl; T22H2.6; Caenorhabditis elegans.
DR   WormBase; WBGene00011936; T22H2.6.
DR   WormPep; T22H2.6b; CE24005.
DR   InterPro; IPR000118; Granulin.
DR   Pfam; PF00396; Granulin; 2.
DR   SMART; SM00277; GRAN; 3.
DR   PROSITE; PS00799; GRANULINS; 2.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   345 AA;  38122 MW;  D93C75167C3650B9 CRC64;

  Query Match             65.6%;  Score 40;  DB 2;  Length 345;
  Best Local Similarity   50.0%;  Pred. No. 50;
  Matches    6;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          3 CXXGPXTWXCXP 14
              |  |  || | |
Db        111 CKLGDNTWGCCP 122


RESULT 15
Q9U362_CAEEL
ID   Q9U362_CAEEL PRELIMINARY;      PRT;   358 AA.
AC   Q9U362; Q9U361;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein T22H2.6a.
GN   ORFNames=T22H2.6, T22H2.6a;
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC   STRAIN=Bristol N2;
RX   MEDLINE=99069613; PubMed=9851916;
RG   The C. elegans sequencing consortium;
RT   "Genome sequence of the nematode C. elegans: a platform for
RT   investigating biology.";
RL   Science 282:2012-2018(1998).
DR   EMBL; Z81595; CAB54304.1; -; Genomic_DNA.
DR   PIR; T25137; T25137.
DR   PIR; T25138; T25138.
DR   HSSP; P28799; 1G26.
DR   Ensembl; T22H2.6; Caenorhabditis elegans.
DR   WormBase; WBGene00011936; T22H2.6.
DR   WormPep; T22H2.6a; CE24004.
DR   InterPro; IPR000118; Granulin.
DR   Pfam; PF00396; Granulin; 3.
DR   SMART; SM00277; GRAN; 3.
DR   PROSITE; PS00799; GRANULINS; 2.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   358 AA;  39754 MW;  2AD5B8F9B70D1595 CRC64;

  Query Match             65.6%;  Score 40;  DB 2;  Length 358;
  Best Local Similarity   50.0%;  Pred. No. 52;
  Matches    6;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;
```

```
Qy          3 CXXGPXTWXCXP 14
              |  |  || | |
Db        111 CKLGDNTWGCCP 122
```

Search completed: March 31, 2006, 16:35:03
Job time : 55.4478 secs

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: March 31, 2006, 16:35:37 ; Search time 13.7214 Seconds
(without alignments)
84.354 Million cell updates/sec

Title: US-10-609-217-83
Perfect score: 61
Sequence: 1 YXCXXGPXTWXCXP 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/6_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/H_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/RE_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	58	95.1	20	1	US-08-484-135-46	Sequence 46, Appl
2	58	95.1	20	1	US-08-484-635-219	Sequence 219, App
3	58	95.1	20	1	US-08-484-631-219	Sequence 219, App
4	58	95.1	20	1	US-08-827-570-219	Sequence 219, App
5	58	95.1	22	1	US-08-484-135-68	Sequence 68, Appl
6	58	95.1	22	1	US-08-484-635-25	Sequence 25, Appl
7	58	95.1	22	1	US-08-484-631-25	Sequence 25, Appl
8	58	95.1	22	1	US-08-827-570-25	Sequence 25, Appl
9	57	93.4	20	1	US-08-484-135-11	Sequence 11, Appl
10	57	93.4	20	1	US-08-484-135-35	Sequence 35, Appl
11	57	93.4	20	1	US-08-484-135-87	Sequence 87, Appl
12	57	93.4	20	1	US-08-484-635-11	Sequence 11, Appl
13	57	93.4	20	1	US-08-484-635-194	Sequence 194, App
14	57	93.4	20	1	US-08-484-635-213	Sequence 213, App
15	57	93.4	20	1	US-08-484-631-11	Sequence 11, Appl
16	57	93.4	20	1	US-08-484-631-194	Sequence 194, App
17	57	93.4	20	1	US-08-484-631-213	Sequence 213, App
18	57	93.4	20	1	US-08-827-570-11	Sequence 11, Appl
19	57	93.4	20	1	US-08-827-570-194	Sequence 194, App
20	57	93.4	20	1	US-08-827-570-213	Sequence 213, App
21	57	93.4	20	2	US-08-905-310-5	Sequence 5, Appli
22	57	93.4	20	2	US-09-428-082B-89	Sequence 89, Appl
23	57	93.4	20	2	US-09-428-082B-1030	Sequence 1030, Ap
24	57	93.4	22	1	US-08-484-135-74	Sequence 74, Appl
25	57	93.4	22	1	US-08-484-635-33	Sequence 33, Appl
26	57	93.4	22	1	US-08-484-631-33	Sequence 33, Appl
27	57	93.4	22	1	US-08-827-570-33	Sequence 33, Appl
28	57	93.4	26	1	US-08-484-635-94	Sequence 94, Appl
29	57	93.4	26	1	US-08-484-635-242	Sequence 242, App
30	57	93.4	26	1	US-08-484-631-94	Sequence 94, Appl
31	57	93.4	26	1	US-08-484-631-242	Sequence 242, App
32	57	93.4	26	1	US-08-827-570-94	Sequence 94, Appl
33	57	93.4	26	1	US-08-827-570-242	Sequence 242, App
34	56	91.8	14	1	US-08-484-635-198	Sequence 198, App
35	56	91.8	14	1	US-08-484-631-198	Sequence 198, App
36	56	91.8	14	1	US-08-827-570-198	Sequence 198, App
37	56	91.8	16	1	US-08-484-635-197	Sequence 197, App
38	56	91.8	16	1	US-08-484-631-197	Sequence 197, App
39	56	91.8	16	1	US-08-827-570-197	Sequence 197, App
40	56	91.8	18	1	US-08-484-135-13	Sequence 13, Appl
41	56	91.8	18	1	US-08-484-635-13	Sequence 13, Appl
42	56	91.8	18	1	US-08-484-635-245	Sequence 245, App
43	56	91.8	18	1	US-08-484-631-13	Sequence 13, Appl
44	56	91.8	18	1	US-08-484-631-245	Sequence 245, App
45	56	91.8	18	1	US-08-827-570-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1
US-08-484-135-46
; Sequence 46, Application US/08484135
; Patent No. 5767078
; GENERAL INFORMATION:
; APPLICANT: Johnson, Dana L
; APPLICANT: Zivin, Robert A
; TITLE OF INVENTION: AGONIST PEPTIDE DIMERS
; NUMBER OF SEQUENCES: 93
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Frank S. DiGiglio
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: U.S.A..
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,135
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: DiGiglio, Frank S
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9594
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-484-135-46

Query Match 95.1%; Score 58; DB 1; Length 20;
Best Local Similarity 57.1%; Pred. No. 0.0041;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
Qy          1 YXCXXGPXTWXCXP 14
              | |  || || | |
Db          4 YSCRMGPTTWVCTP 17
```

```
RESULT 2
US-08-484-635-219
; Sequence 219, Application US/08484635
; Patent No. 5773569
;  GENERAL INFORMATION:
;    APPLICANT:  Wrighton, Nicholas C.
;    APPLICANT:  Dower, William J.
;    APPLICANT:  Chang, Ray S.
;    APPLICANT:  Kashyap, Arun K.
;    APPLICANT:  Jolliffe, Linda K.
;    APPLICANT:  Johnson, Dana
;    APPLICANT:  Mulcahy, Linda
;    TITLE OF INVENTION:  Compounds and Peptides That Bind to the
;    TITLE OF INVENTION:  Erythropoietin Receptor
;    NUMBER OF SEQUENCES:  259
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew
;      STREET:  One Market Plaza, Steuart Street Tower
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94105-1492
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/484,635
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/155,940
;      FILING DATE:  19-NOV-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Garrett-Wackowski, Eugenia
;      REGISTRATION NUMBER:  37,330
;      REFERENCE/DOCKET NUMBER:  16528A-43-1-1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 543-9600
;      TELEFAX:  (415) 543-5043
;  INFORMATION FOR SEQ ID NO:  219:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-484-635-219

  Query Match             95.1%;  Score 58;  DB 1;  Length 20;
  Best Local Similarity   57.1%;  Pred. No. 0.0041;
  Matches    8;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 YXCXXGPXTWXCXP 14
              | |  || || | |
Db          4 YSCRMGPTTWVCTP 17


RESULT 3
US-08-484-631-219
; Sequence 219, Application US/08484631
; Patent No. 5830851
;  GENERAL INFORMATION:
;    APPLICANT:  Wrighton, Nicholas C.
;    APPLICANT:  Dower, William J.
;    APPLICANT:  Chang, Ray S.
;    APPLICANT:  Kashyap, Arun K.
;    APPLICANT:  Jolliffe, Linda K.
;    APPLICANT:  Johnson, Dana
;    APPLICANT:  Mulcahy, Linda
;    TITLE OF INVENTION:  Compounds and Peptides That Bind to the
;    TITLE OF INVENTION:  Erythropoietin Receptor
;    NUMBER OF SEQUENCES:  259
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew
;      STREET:  One Market Plaza, Steuart Street Tower
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94105-1492
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/484,631
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/155,940
;      FILING DATE:  19-NOV-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Garrett-Wackowski, Eugenia
;      REGISTRATION NUMBER:  37,330
;      REFERENCE/DOCKET NUMBER:  16528A-43-1-2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 543-9600
;      TELEFAX:  (415) 543-5043
;  INFORMATION FOR SEQ ID NO:  219:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-484-631-219

  Query Match             95.1%;  Score 58;  DB 1;  Length 20;
  Best Local Similarity   57.1%;  Pred. No. 0.0041;
  Matches    8;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 YXCXXGPXTWXCXP 14
              | |  || || | |
Db          4 YSCRMGPTTWVCTP 17


RESULT 4
US-08-827-570-219
; Sequence 219, Application US/08827570
; Patent No. 5986047
;  GENERAL INFORMATION:
;    APPLICANT:  Wrighton, Nicholas C.
;    APPLICANT:  Dower, William J.
;    APPLICANT:  Chang, Ray S.
;    APPLICANT:  Kashyap, Arun K.
;    APPLICANT:  Jolliffe, Linda K.
;    APPLICANT:  Johnson, Dana
;    APPLICANT:  Mulcahy, Linda
;    TITLE OF INVENTION:  Compounds and Peptides That Bind to the
;    TITLE OF INVENTION:  Erythropoietin Receptor
;    NUMBER OF SEQUENCES:  259
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew
;      STREET:  One Market Plaza, Steuart Street Tower
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94105-1492
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
```

```
;      APPLICATION NUMBER:  US/08/827,570
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/484,635
;      FILING DATE:  07-JUN-1995
;      APPLICATION NUMBER:  US 08/155,940
;      FILING DATE:  19-NOV-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Garrett-Wackowski, Eugenia
;      REGISTRATION NUMBER:  37,330
;      REFERENCE/DOCKET NUMBER:  16528A-43-1-1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 543-9600
;      TELEFAX:  (415) 543-5043
;  INFORMATION FOR SEQ ID NO:  219:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-827-570-219

  Query Match             95.1%;  Score 58;  DB 1;  Length 20;
  Best Local Similarity   57.1%;  Pred. No. 0.0041;
  Matches    8;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 YXCXXGPXTWXCXP 14
              | |  || || | |
Db          4 YSCRMGPTTWVCTP 17


RESULT 5
US-08-484-135-68
; Sequence 68, Application US/08484135
; Patent No. 5767078
;  GENERAL INFORMATION:
;    APPLICANT:  Johnson, Dana L
;    APPLICANT:  Zivin, Robert A
;    TITLE OF INVENTION:  AGONIST PEPTIDE DIMERS
;    NUMBER OF SEQUENCES:  93
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Frank S. DiGiglio
;      STREET:  400 Garden City Plaza
;      CITY:  Garden City
;      STATE:  New York
;      COUNTRY:  U.S.A..
;      ZIP:  11530
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/484,135
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DiGiglio, Frank S
;      REGISTRATION NUMBER:  31,346
;      REFERENCE/DOCKET NUMBER:  9594
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (516) 742-4343
;      TELEFAX:  (516) 742-4366
;  INFORMATION FOR SEQ ID NO:  68:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  22 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-484-135-68

  Query Match             95.1%;  Score 58;  DB 1;  Length 22;
  Best Local Similarity   57.1%;  Pred. No. 0.0044;
  Matches    8;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 YXCXXGPXTWXCXP 14
              | |  || || | |
Db          4 YSCFMGPTTWVCSP 17


RESULT 6
US-08-484-635-25
; Sequence 25, Application US/08484635
; Patent No. 5773569
;  GENERAL INFORMATION:
;    APPLICANT:  Wrighton, Nicholas C.
;    APPLICANT:  Dower, William J.
;    APPLICANT:  Chang, Ray S.
;    APPLICANT:  Kashyap, Arun K.
;    APPLICANT:  Jolliffe, Linda K.
;    APPLICANT:  Johnson, Dana
;    APPLICANT:  Mulcahy, Linda
;    TITLE OF INVENTION:  Compounds and Peptides That Bind to the
;    TITLE OF INVENTION:  Erythropoietin Receptor
;    NUMBER OF SEQUENCES:  259
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew
;      STREET:  One Market Plaza, Steuart Street Tower
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94105-1492
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/484,635
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/155,940
;      FILING DATE:  19-NOV-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Garrett-Wackowski, Eugenia
;      REGISTRATION NUMBER:  37,330
;      REFERENCE/DOCKET NUMBER:  16528A-43-1-1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 543-9600
;      TELEFAX:  (415) 543-5043
;  INFORMATION FOR SEQ ID NO:  25:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  22 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-484-635-25

  Query Match             95.1%;  Score 58;  DB 1;  Length 22;
  Best Local Similarity   57.1%;  Pred. No. 0.0044;
  Matches    8;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 YXCXXGPXTWXCXP 14
              | |  || || | |
Db          4 YSCFMGPTTWVCSP 17


RESULT 7
US-08-484-631-25
; Sequence 25, Application US/08484631
; Patent No. 5830851
```

```
;  GENERAL INFORMATION:
;    APPLICANT:  Wrighton, Nicholas C.
;    APPLICANT:  Dower, William J.
;    APPLICANT:  Chang, Ray S.
;    APPLICANT:  Kashyap, Arun K.
;    APPLICANT:  Jolliffe, Linda K.
;    APPLICANT:  Johnson, Dana
;    APPLICANT:  Mulcahy, Linda
;    TITLE OF INVENTION:  Compounds and Peptides That Bind to the
;    TITLE OF INVENTION:  Erythropoietin Receptor
;    NUMBER OF SEQUENCES:  259
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew
;      STREET:  One Market Plaza, Steuart Street Tower
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94105-1492
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/484,631
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/155,940
;      FILING DATE:  19-NOV-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Garrett-Wackowski, Eugenia
;      REGISTRATION NUMBER:  37,330
;      REFERENCE/DOCKET NUMBER:  16528A-43-1-2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 543-9600
;      TELEFAX:  (415) 543-5043
;  INFORMATION FOR SEQ ID NO:  25:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  22 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-484-631-25

  Query Match             95.1%;  Score 58;  DB 1;  Length 22;
  Best Local Similarity   57.1%;  Pred. No. 0.0044;
  Matches    8;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 YXCXXGPXTWXCXP 14
              | |  || || | |
Db          4 YSCFMGPTTWVCSP 17


RESULT 8
US-08-827-570-25
; Sequence 25, Application US/08827570
; Patent No. 5986047
;  GENERAL INFORMATION:
;    APPLICANT:  Wrighton, Nicholas C.
;    APPLICANT:  Dower, William J.
;    APPLICANT:  Chang, Ray S.
;    APPLICANT:  Kashyap, Arun K.
;    APPLICANT:  Jolliffe, Linda K.
;    APPLICANT:  Johnson, Dana
;    APPLICANT:  Mulcahy, Linda
;    TITLE OF INVENTION:  Compounds and Peptides That Bind to the
;    TITLE OF INVENTION:  Erythropoietin Receptor
;    NUMBER OF SEQUENCES:  259
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew
;      STREET:  One Market Plaza, Steuart Street Tower
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94105-1492
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/827,570
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/484,635
;      FILING DATE:  07-JUN-1995
;      APPLICATION NUMBER:  US 08/155,940
;      FILING DATE:  19-NOV-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Garrett-Wackowski, Eugenia
;      REGISTRATION NUMBER:  37,330
;      REFERENCE/DOCKET NUMBER:  16528A-43-1-1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 543-9600
;      TELEFAX:  (415) 543-5043
;  INFORMATION FOR SEQ ID NO:  25:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  22 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-827-570-25

  Query Match             95.1%;  Score 58;  DB 1;  Length 22;
  Best Local Similarity   57.1%;  Pred. No. 0.0044;
  Matches    8;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 YXCXXGPXTWXCXP 14
              | |  || || | |
Db          4 YSCFMGPTTWVCSP 17


RESULT 9
US-08-484-135-11
; Sequence 11, Application US/08484135
; Patent No. 5767078
;  GENERAL INFORMATION:
;    APPLICANT:  Johnson, Dana L
;    APPLICANT:  Zivin, Robert A
;    TITLE OF INVENTION:  AGONIST PEPTIDE DIMERS
;    NUMBER OF SEQUENCES:  93
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Frank S. DiGiglio
;      STREET:  400 Garden City Plaza
;      CITY:  Garden City
;      STATE:  New York
;      COUNTRY:  U.S.A..
;      ZIP:  11530
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/484,135
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DiGiglio, Frank S
;      REGISTRATION NUMBER:  31,346
;      REFERENCE/DOCKET NUMBER:  9594
;    TELECOMMUNICATION INFORMATION:
```

```
;      TELEPHONE:  (516) 742-4343
;      TELEFAX:  (516) 742-4366
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-484-135-11

  Query Match             93.4%;  Score 57;  DB 1;  Length 20;
  Best Local Similarity   57.1%;  Pred. No. 0.0059;
  Matches    8;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 YXCXXGPXTWXCXP 14
              | |  || || | |
Db          4 YACRMGPITWVCSP 17


RESULT 10
US-08-484-135-35
; Sequence 35, Application US/08484135
; Patent No. 5767078
;  GENERAL INFORMATION:
;    APPLICANT:  Johnson, Dana L
;    APPLICANT:  Zivin, Robert A
;    TITLE OF INVENTION:  AGONIST PEPTIDE DIMERS
;    NUMBER OF SEQUENCES:  93
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Frank S. DiGiglio
;      STREET:  400 Garden City Plaza
;      CITY:  Garden City
;      STATE:  New York
;      COUNTRY:  U.S.A..
;      ZIP:  11530
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/484,135
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DiGiglio, Frank S
;      REGISTRATION NUMBER:  31,346
;      REFERENCE/DOCKET NUMBER:  9594
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (516) 742-4343
;      TELEFAX:  (516) 742-4366
;  INFORMATION FOR SEQ ID NO:  35:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-484-135-35

  Query Match             93.4%;  Score 57;  DB 1;  Length 20;
  Best Local Similarity   57.1%;  Pred. No. 0.0059;
  Matches    8;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 YXCXXGPXTWXCXP 14
              | |  || || | |
Db          4 YSCRMGPMTWVCSP 17


RESULT 11
US-08-484-135-87
; Sequence 87, Application US/08484135
; Patent No. 5767078
;  GENERAL INFORMATION:
;    APPLICANT:  Johnson, Dana L
;    APPLICANT:  Zivin, Robert A
;    TITLE OF INVENTION:  AGONIST PEPTIDE DIMERS
;    NUMBER OF SEQUENCES:  93
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Frank S. DiGiglio
;      STREET:  400 Garden City Plaza
;      CITY:  Garden City
;      STATE:  New York
;      COUNTRY:  U.S.A..
;      ZIP:  11530
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/484,135
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DiGiglio, Frank S
;      REGISTRATION NUMBER:  31,346
;      REFERENCE/DOCKET NUMBER:  9594
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (516) 742-4343
;      TELEFAX:  (516) 742-4366
;  INFORMATION FOR SEQ ID NO:  87:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-484-135-87

  Query Match             93.4%;  Score 57;  DB 1;  Length 20;
  Best Local Similarity   57.1%;  Pred. No. 0.0059;
  Matches    8;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 YXCXXGPXTWXCXP 14
              | |  || || | |
Db          4 YACRMGPITWVCSP 17


RESULT 12
US-08-484-635-11
; Sequence 11, Application US/08484635
; Patent No. 5773569
;  GENERAL INFORMATION:
;    APPLICANT:  Wrighton, Nicholas C.
;    APPLICANT:  Dower, William J.
;    APPLICANT:  Chang, Ray S.
;    APPLICANT:  Kashyap, Arun K.
;    APPLICANT:  Jolliffe, Linda K.
;    APPLICANT:  Johnson, Dana
;    APPLICANT:  Mulcahy, Linda
;    TITLE OF INVENTION:  Compounds and Peptides That Bind to the
;    TITLE OF INVENTION:  Erythropoietin Receptor
;    NUMBER OF SEQUENCES:  259
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew
;      STREET:  One Market Plaza, Steuart Street Tower
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94105-1492
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
```

```
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/484,635
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/155,940
;      FILING DATE:  19-NOV-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Garrett-Wackowski, Eugenia
;      REGISTRATION NUMBER:  37,330
;      REFERENCE/DOCKET NUMBER:  16528A-43-1-1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 543-9600
;      TELEFAX:  (415) 543-5043
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-484-635-11

  Query Match             93.4%;  Score 57;  DB 1;  Length 20;
  Best Local Similarity   57.1%;  Pred. No. 0.0059;
  Matches    8;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 YXCXXGPXTWXCXP 14
              | |  || || | |
Db          4 YACRMGPITWVCSP 17


RESULT 13
US-08-484-635-194
; Sequence 194, Application US/08484635
; Patent No. 5773569
;  GENERAL INFORMATION:
;    APPLICANT:  Wrighton, Nicholas C.
;    APPLICANT:  Dower, William J.
;    APPLICANT:  Chang, Ray S.
;    APPLICANT:  Kashyap, Arun K.
;    APPLICANT:  Jolliffe, Linda K.
;    APPLICANT:  Johnson, Dana
;    APPLICANT:  Mulcahy, Linda
;    TITLE OF INVENTION:  Compounds and Peptides That Bind to the
;    TITLE OF INVENTION:  Erythropoietin Receptor
;    NUMBER OF SEQUENCES:  259
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew
;      STREET:  One Market Plaza, Steuart Street Tower
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94105-1492
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/484,635
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/155,940
;      FILING DATE:  19-NOV-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Garrett-Wackowski, Eugenia
;      REGISTRATION NUMBER:  37,330
;      REFERENCE/DOCKET NUMBER:  16528A-43-1-1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 543-9600
;      TELEFAX:  (415) 543-5043
;  INFORMATION FOR SEQ ID NO:  194:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-484-635-194

  Query Match             93.4%;  Score 57;  DB 1;  Length 20;
  Best Local Similarity   57.1%;  Pred. No. 0.0059;
  Matches    8;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 YXCXXGPXTWXCXP 14
              | |  || || | |
Db          4 YSCHFGPATWVCKP 17


RESULT 14
US-08-484-635-213
; Sequence 213, Application US/08484635
; Patent No. 5773569
;  GENERAL INFORMATION:
;    APPLICANT:  Wrighton, Nicholas C.
;    APPLICANT:  Dower, William J.
;    APPLICANT:  Chang, Ray S.
;    APPLICANT:  Kashyap, Arun K.
;    APPLICANT:  Jolliffe, Linda K.
;    APPLICANT:  Johnson, Dana
;    APPLICANT:  Mulcahy, Linda
;    TITLE OF INVENTION:  Compounds and Peptides That Bind to the
;    TITLE OF INVENTION:  Erythropoietin Receptor
;    NUMBER OF SEQUENCES:  259
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew
;      STREET:  One Market Plaza, Steuart Street Tower
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94105-1492
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/484,635
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/155,940
;      FILING DATE:  19-NOV-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Garrett-Wackowski, Eugenia
;      REGISTRATION NUMBER:  37,330
;      REFERENCE/DOCKET NUMBER:  16528A-43-1-1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 543-9600
;      TELEFAX:  (415) 543-5043
;  INFORMATION FOR SEQ ID NO:  213:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-484-635-213

  Query Match             93.4%;  Score 57;  DB 1;  Length 20;
  Best Local Similarity   57.1%;  Pred. No. 0.0059;
  Matches    8;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;
```

```
Qy          1 YXCXXGPXTWXCXP 14
              | |  || || | |
Db          4 YSCRMGPMTWVCSP 17


RESULT 15
US-08-484-631-11
; Sequence 11, Application US/08484631
; Patent No. 5830851
;  GENERAL INFORMATION:
;    APPLICANT:  Wrighton, Nicholas C.
;    APPLICANT:  Dower, William J.
;    APPLICANT:  Chang, Ray S.
;    APPLICANT:  Kashyap, Arun K.
;    APPLICANT:  Jolliffe, Linda K.
;    APPLICANT:  Johnson, Dana
;    APPLICANT:  Mulcahy, Linda
;    TITLE OF INVENTION:  Compounds and Peptides That Bind to the
;    TITLE OF INVENTION:  Erythropoietin Receptor
;    NUMBER OF SEQUENCES:  259
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew
;      STREET:  One Market Plaza, Steuart Street Tower
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94105-1492
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/484,631
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/155,940
;      FILING DATE:  19-NOV-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Garrett-Wackowski, Eugenia
;      REGISTRATION NUMBER:  37,330
;      REFERENCE/DOCKET NUMBER:  16528A-43-1-2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 543-9600
;      TELEFAX:  (415) 543-5043
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-484-631-11

  Query Match             93.4%;  Score 57;  DB 1;  Length 20;
  Best Local Similarity   57.1%;  Pred. No. 0.0059;
  Matches    8;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 YXCXXGPXTWXCXP 14
              | |  || || | |
Db          4 YACRMGPITWVCSP 17


Search completed: March 31, 2006, 16:40:34
Job time : 13.7214 secs
```

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: March 31, 2006, 17:21:12 ; Search time 43.1144 Seconds
(without alignments)
135.676 Million cell updates/sec

Title: US-10-609-217-83
Perfect score: 61
Sequence: 1 YXCXXGPXTWXCXP 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA_Main:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	57	93.4	20	3	US-09-863-600E-11	Sequence 11, Appl
2	57	93.4	20	3	US-09-863-600E-32	Sequence 32, Appl
3	57	93.4	20	4	US-10-609-217-89	Sequence 89, Appl
4	57	93.4	20	4	US-10-609-217-1030	Sequence 1030, Ap
5	57	93.4	20	4	US-10-632-388-89	Sequence 89, Appl
6	57	93.4	20	4	US-10-632-388-1030	Sequence 1030, Ap
7	57	93.4	20	4	US-10-651-723-89	Sequence 89, Appl
8	57	93.4	20	4	US-10-651-723-1030	Sequence 1030, Ap
9	57	93.4	20	4	US-10-645-761-89	Sequence 89, Appl
10	57	93.4	20	4	US-10-645-761-1030	Sequence 1030, Ap
11	57	93.4	20	4	US-10-666-696-89	Sequence 89, Appl
12	57	93.4	20	4	US-10-666-696-1030	Sequence 1030, Ap
13	57	93.4	20	4	US-10-653-048-89	Sequence 89, Appl
14	57	93.4	20	4	US-10-653-048-1030	Sequence 1030, Ap
15	57	93.4	20	4	US-10-460-550-9	Sequence 9, Appli
16	57	93.4	20	5	US-10-810-362-5	Sequence 5, Appli
17	57	93.4	20	5	US-10-645-784-89	Sequence 89, Appl
18	57	93.4	20	5	US-10-645-784-1030	Sequence 1030, Ap
19	57	93.4	20	5	US-10-609-783B-7	Sequence 7, Appli
20	57	93.4	20	5	US-10-609-783B-34	Sequence 34, Appl
21	57	93.4	20	5	US-10-460-550-9	Sequence 9, Appli
22	56	91.8	14	3	US-09-863-600E-24	Sequence 24, Appl
23	56	91.8	16	3	US-09-863-600E-23	Sequence 23, Appl
24	56	91.8	18	3	US-09-863-600E-13	Sequence 13, Appl
25	56	91.8	18	3	US-09-863-600E-22	Sequence 22, Appl
26	56	91.8	18	4	US-10-609-217-425	Sequence 425, App
27	56	91.8	18	4	US-10-609-217-1036	Sequence 1036, Ap
28	56	91.8	18	4	US-10-632-388-425	Sequence 425, App
29	56	91.8	18	4	US-10-632-388-1036	Sequence 1036, Ap
30	56	91.8	18	4	US-10-651-723-425	Sequence 425, App
31	56	91.8	18	4	US-10-651-723-1036	Sequence 1036, Ap
32	56	91.8	18	4	US-10-645-761-425	Sequence 425, App
33	56	91.8	18	4	US-10-645-761-1036	Sequence 1036, Ap
34	56	91.8	18	4	US-10-666-696-425	Sequence 425, App
35	56	91.8	18	4	US-10-666-696-1036	Sequence 1036, Ap
36	56	91.8	18	4	US-10-653-048-425	Sequence 425, App
37	56	91.8	18	4	US-10-653-048-1036	Sequence 1036, Ap
38	56	91.8	18	5	US-10-645-784-425	Sequence 425, App
39	56	91.8	18	5	US-10-645-784-1036	Sequence 1036, Ap
40	56	91.8	18	5	US-10-609-783B-21	Sequence 21, Appl
41	56	91.8	18	5	US-10-609-783B-36	Sequence 36, Appl
42	56	91.8	20	3	US-09-858-935B-66	Sequence 66, Appl
43	56	91.8	20	3	US-09-863-600E-8	Sequence 8, Appli
44	56	91.8	20	3	US-09-863-600E-14	Sequence 14, Appl
45	56	91.8	20	3	US-09-863-600E-43	Sequence 43, Appl

ALIGNMENTS

RESULT 1
US-09-863-600E-11
; Sequence 11, Application US/09863600E
; Publication No. US20030130197A1
; GENERAL INFORMATION:
; APPLICANT: Smith-Swintosky, Virginia
; APPLICANT: Renzi, Michael
; APPLICANT: Plata-Salaman, Carlos
; APPLICANT: Jolliffe, Linda
; APPLICANT: Farrell, Francis
; APPLICANT: Johnson, Dana
; TITLE OF INVENTION: Neuroprotective Peptides
; FILE REFERENCE: PRI-0014 (ORT-1436)
; CURRENT APPLICATION NUMBER: US/09/863,600E
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: 60/207,654
; PRIOR FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 11
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide
US-09-863-600E-11

Query Match 93.4%; Score 57; DB 3; Length 20;
Best Local Similarity 57.1%; Pred. No. 0.012;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
Qy          1 YXCXXGPXTWXCXP 14
              | |  || || | |
Db          4 YACRMGPITWVCSP 17
```

RESULT 2
US-09-863-600E-32
; Sequence 32, Application US/09863600E
; Publication No. US20030130197A1
; GENERAL INFORMATION:
; APPLICANT: Smith-Swintosky, Virginia
; APPLICANT: Renzi, Michael
; APPLICANT: Plata-Salaman, Carlos
; APPLICANT: Jolliffe, Linda
; APPLICANT: Farrell, Francis
; APPLICANT: Johnson, Dana
; TITLE OF INVENTION: Neuroprotective Peptides
; FILE REFERENCE: PRI-0014 (ORT-1436)
; CURRENT APPLICATION NUMBER: US/09/863,600E

; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: 60/207,654
; PRIOR FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 32
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide
US-09-863-600E-32

Query Match 93.4%; Score 57; DB 3; Length 20;
Best Local Similarity 57.1%; Pred. No. 0.012;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 YXCXXGPXTWXCXP 14
| | || || | |
Db 4 YSCHFGPATWVCKP 17

RESULT 3
US-10-609-217-89
; Sequence 89, Application US/10609217
; Publication No. US20040044188A1
; GENERAL INFORMATION:
; APPLICANT: FEIGE, ULRICH
; APPLICANT: LIU, CHUAN-FA
; APPLICANT: CHEETHAM, JANET C.
; APPLICANT: BOONE, THOMAS CHARLES
; TITLE OF INVENTION: MODIFIED PEPTIDES AS THERAPEUTIC AGENTS
; FILE REFERENCE: A-527
; CURRENT APPLICATION NUMBER: US/10/609,217
; CURRENT FILING DATE: 2003-06-27
; PRIOR APPLICATION NUMBER: US/09/428,082B
; PRIOR FILING DATE: 1999-10-22
; PRIOR APPLICATION NUMBER: 60/105,371
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 1133
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 89
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: EPO-MIMETIC PEPTIDE
US-10-609-217-89

Query Match 93.4%; Score 57; DB 4; Length 20;
Best Local Similarity 57.1%; Pred. No. 0.012;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 YXCXXGPXTWXCXP 14
| | || || | |
Db 4 YACRMGPITWVCSP 17

RESULT 4
US-10-609-217-1030
; Sequence 1030, Application US/10609217
; Publication No. US20040044188A1
; GENERAL INFORMATION:
; APPLICANT: FEIGE, ULRICH
; APPLICANT: LIU, CHUAN-FA
; APPLICANT: CHEETHAM, JANET C.
; APPLICANT: BOONE, THOMAS CHARLES
; TITLE OF INVENTION: MODIFIED PEPTIDES AS THERAPEUTIC AGENTS
; FILE REFERENCE: A-527
; CURRENT APPLICATION NUMBER: US/10/609,217
; CURRENT FILING DATE: 2003-06-27
; PRIOR APPLICATION NUMBER: US/09/428,082B
; PRIOR FILING DATE: 1999-10-22
; PRIOR APPLICATION NUMBER: 60/105,371
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 1133
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1030
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: EPO-MIMETIC PEPTIDE
US-10-609-217-1030

Query Match 93.4%; Score 57; DB 4; Length 20;
Best Local Similarity 57.1%; Pred. No. 0.012;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 YXCXXGPXTWXCXP 14
| | || || | |
Db 4 YACRMGPITWVCSP 17

RESULT 5
US-10-632-388-89
; Sequence 89, Application US/10632388
; Publication No. US20040053845A1
; GENERAL INFORMATION:
; APPLICANT: FEIGE, ULRICH
; APPLICANT: LIU, CHUAN-FA
; APPLICANT: CHEETHAM, JANET C.
; APPLICANT: BOONE, THOMAS CHARLES
; TITLE OF INVENTION: MODIFIED PEPTIDES AS THERAPEUTIC AGENTS
; FILE REFERENCE: A-527
; CURRENT APPLICATION NUMBER: US/10/632,388
; CURRENT FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: US/09/428,082B
; PRIOR FILING DATE: 1999-10-22
; PRIOR APPLICATION NUMBER: 60/105,371
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 1133
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 89
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: EPO-MIMETIC PEPTIDE
US-10-632-388-89

Query Match 93.4%; Score 57; DB 4; Length 20;
Best Local Similarity 57.1%; Pred. No. 0.012;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 YXCXXGPXTWXCXP 14
| | || || | |
Db 4 YACRMGPITWVCSP 17

RESULT 6
US-10-632-388-1030
; Sequence 1030, Application US/10632388
; Publication No. US20040053845A1
; GENERAL INFORMATION:
; APPLICANT: FEIGE, ULRICH
; APPLICANT: LIU, CHUAN-FA
; APPLICANT: CHEETHAM, JANET C.
; APPLICANT: BOONE, THOMAS CHARLES
; TITLE OF INVENTION: MODIFIED PEPTIDES AS THERAPEUTIC AGENTS
; FILE REFERENCE: A-527
; CURRENT APPLICATION NUMBER: US/10/632,388
; CURRENT FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: US/09/428,082B
; PRIOR FILING DATE: 1999-10-22
; PRIOR APPLICATION NUMBER: 60/105,371

; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 1133
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1030
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: EPO-MIMETIC PEPTIDE
US-10-632-388-1030

Query Match 93.4%; Score 57; DB 4; Length 20;
Best Local Similarity 57.1%; Pred. No. 0.012;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 YXCXXGPXTWXCXP 14
| | || || | |
Db 4 YACRMGPITWVCSP 17

RESULT 7
US-10-651-723-89
; Sequence 89, Application US/10651723
; Publication No. US20040057953A1
; GENERAL INFORMATION:
; APPLICANT: FEIGE, ULRICH
; APPLICANT: LIU, CHUAN-FA
; APPLICANT: CHEETHAM, JANET C.
; APPLICANT: BOONE, THOMAS CHARLES
; TITLE OF INVENTION: MODIFIED PEPTIDES AS THERAPEUTIC AGENTS
; FILE REFERENCE: A-527
; CURRENT APPLICATION NUMBER: US/10/651,723
; CURRENT FILING DATE: 2003-08-29
; PRIOR APPLICATION NUMBER: US/09/428,082B
; PRIOR FILING DATE: 1999-10-22
; PRIOR APPLICATION NUMBER: 60/105,371
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 1133
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 89
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: EPO-MIMETIC PEPTIDE
US-10-651-723-89

Query Match 93.4%; Score 57; DB 4; Length 20;
Best Local Similarity 57.1%; Pred. No. 0.012;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 YXCXXGPXTWXCXP 14
| | || || | |
Db 4 YACRMGPITWVCSP 17

RESULT 8
US-10-651-723-1030
; Sequence 1030, Application US/10651723
; Publication No. US20040057953A1
; GENERAL INFORMATION:
; APPLICANT: FEIGE, ULRICH
; APPLICANT: LIU, CHUAN-FA
; APPLICANT: CHEETHAM, JANET C.
; APPLICANT: BOONE, THOMAS CHARLES
; TITLE OF INVENTION: MODIFIED PEPTIDES AS THERAPEUTIC AGENTS
; FILE REFERENCE: A-527
; CURRENT APPLICATION NUMBER: US/10/651,723
; CURRENT FILING DATE: 2003-08-29
; PRIOR APPLICATION NUMBER: US/09/428,082B
; PRIOR FILING DATE: 1999-10-22
; PRIOR APPLICATION NUMBER: 60/105,371
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 1133
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1030
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: EPO-MIMETIC PEPTIDE
US-10-651-723-1030

Query Match 93.4%; Score 57; DB 4; Length 20;
Best Local Similarity 57.1%; Pred. No. 0.012;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 YXCXXGPXTWXCXP 14
| | || || | |
Db 4 YACRMGPITWVCSP 17

RESULT 9
US-10-645-761-89
; Sequence 89, Application US/10645761
; Publication No. US20040071712A1
; GENERAL INFORMATION:
; APPLICANT: FEIGE, ULRICH
; APPLICANT: LIU, CHUAN-FA
; APPLICANT: CHEETHAM, JANET C.
; APPLICANT: BOONE, THOMAS CHARLES
; TITLE OF INVENTION: MODIFIED PEPTIDES AS THERAPEUTIC AGENTS
; FILE REFERENCE: A-527
; CURRENT APPLICATION NUMBER: US/10/645,761
; CURRENT FILING DATE: 2003-08-18
; PRIOR APPLICATION NUMBER: US/09/428,082B
; PRIOR FILING DATE: 1999-10-22
; PRIOR APPLICATION NUMBER: 60/105,371
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 1133
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 89
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: EPO-MIMETIC PEPTIDE
US-10-645-761-89

Query Match 93.4%; Score 57; DB 4; Length 20;
Best Local Similarity 57.1%; Pred. No. 0.012;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 YXCXXGPXTWXCXP 14
| | || || | |
Db 4 YACRMGPITWVCSP 17

RESULT 10
US-10-645-761-1030
; Sequence 1030, Application US/10645761
; Publication No. US20040071712A1
; GENERAL INFORMATION:
; APPLICANT: FEIGE, ULRICH
; APPLICANT: LIU, CHUAN-FA
; APPLICANT: CHEETHAM, JANET C.
; APPLICANT: BOONE, THOMAS CHARLES
; TITLE OF INVENTION: MODIFIED PEPTIDES AS THERAPEUTIC AGENTS
; FILE REFERENCE: A-527
; CURRENT APPLICATION NUMBER: US/10/645,761
; CURRENT FILING DATE: 2003-08-18
; PRIOR APPLICATION NUMBER: US/09/428,082B
; PRIOR FILING DATE: 1999-10-22
; PRIOR APPLICATION NUMBER: 60/105,371
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 1133

```
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1030
;   LENGTH: 20
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: EPO-MIMETIC PEPTIDE
US-10-645-761-1030

  Query Match             93.4%;  Score 57;  DB 4;  Length 20;
  Best Local Similarity   57.1%;  Pred. No. 0.012;
  Matches    8;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 YXCXXGPXTWXCXP 14
              | |  || || | |
Db          4 YACRMGPITWVCSP 17


RESULT 11
US-10-666-696-89
; Sequence 89, Application US/10666696
; Publication No. US20040077022A1
; GENERAL INFORMATION:
;  APPLICANT: FEIGE, ULRICH
;  APPLICANT:  LIU, CHUAN-FA
;  APPLICANT:  CHEETHAM, JANET C.
;  APPLICANT:  BOONE, THOMAS CHARLES
;  APPLICANT:  GUDAS, JEAN MARIE
;  TITLE OF INVENTION: MODIFIED PEPTIDES AS THERAPEUTIC AGENTS
;  FILE REFERENCE: A-527A
;  CURRENT APPLICATION NUMBER: US/10/666,696
;  CURRENT FILING DATE:  2003-09-19
;  PRIOR APPLICATION NUMBER: US/09/563,286C
;  PRIOR FILING DATE: 2000-05-03
;  PRIOR APPLICATION NUMBER: 09/428,082
;  PRIOR FILING DATE: 1999-10-22
;  PRIOR APPLICATION NUMBER: 60/105,371
;  PRIOR FILING DATE: 1998-10-23
;  NUMBER OF SEQ ID NOS: 1157
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 89
;   LENGTH: 20
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: EPO-mimetic peptide
US-10-666-696-89

  Query Match             93.4%;  Score 57;  DB 4;  Length 20;
  Best Local Similarity   57.1%;  Pred. No. 0.012;
  Matches    8;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 YXCXXGPXTWXCXP 14
              | |  || || | |
Db          4 YACRMGPITWVCSP 17


RESULT 12
US-10-666-696-1030
; Sequence 1030, Application US/10666696
; Publication No. US20040077022A1
; GENERAL INFORMATION:
;  APPLICANT: FEIGE, ULRICH
;  APPLICANT:  LIU, CHUAN-FA
;  APPLICANT:  CHEETHAM, JANET C.
;  APPLICANT:  BOONE, THOMAS CHARLES
;  APPLICANT:  GUDAS, JEAN MARIE
;  TITLE OF INVENTION: MODIFIED PEPTIDES AS THERAPEUTIC AGENTS
;  FILE REFERENCE: A-527A
;  CURRENT APPLICATION NUMBER: US/10/666,696
;  CURRENT FILING DATE:  2003-09-19
;  PRIOR APPLICATION NUMBER: US/09/563,286C
;  PRIOR FILING DATE: 2000-05-03
;  PRIOR APPLICATION NUMBER: 09/428,082
;  PRIOR FILING DATE: 1999-10-22
;  PRIOR APPLICATION NUMBER: 60/105,371
;  PRIOR FILING DATE: 1998-10-23
;  NUMBER OF SEQ ID NOS: 1157
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1030
;   LENGTH: 20
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: EPO MIMETIC PEPTIDE
US-10-666-696-1030

  Query Match             93.4%;  Score 57;  DB 4;  Length 20;
  Best Local Similarity   57.1%;  Pred. No. 0.012;
  Matches    8;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 YXCXXGPXTWXCXP 14
              | |  || || | |
Db          4 YACRMGPITWVCSP 17


RESULT 13
US-10-653-048-89
; Sequence 89, Application US/10653048
; Publication No. US20040087778A1
; GENERAL INFORMATION:
;  APPLICANT: FEIGE, ULRICH
;  APPLICANT:  LIU, CHUAN-FA
;  APPLICANT:  CHEETHAM, JANET C.
;  APPLICANT:  BOONE, THOMAS CHARLES
;  TITLE OF INVENTION: MODIFIED PEPTIDES AS THERAPEUTIC AGENTS
;  FILE REFERENCE: A-527
;  CURRENT APPLICATION NUMBER: US/10/653,048
;  CURRENT FILING DATE:  2003-08-29
;  PRIOR APPLICATION NUMBER: US/09/428,082B
;  PRIOR FILING DATE: 1999-10-22
;  PRIOR APPLICATION NUMBER: 60/105,371
;  PRIOR FILING DATE: 1998-10-23
;  NUMBER OF SEQ ID NOS: 1133
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 89
;   LENGTH: 20
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: EPO-MIMETIC PEPTIDE
US-10-653-048-89

  Query Match             93.4%;  Score 57;  DB 4;  Length 20;
  Best Local Similarity   57.1%;  Pred. No. 0.012;
  Matches    8;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 YXCXXGPXTWXCXP 14
              | |  || || | |
Db          4 YACRMGPITWVCSP 17


RESULT 14
US-10-653-048-1030
; Sequence 1030, Application US/10653048
; Publication No. US20040087778A1
; GENERAL INFORMATION:
;  APPLICANT: FEIGE, ULRICH
;  APPLICANT:  LIU, CHUAN-FA
;  APPLICANT:  CHEETHAM, JANET C.
;  APPLICANT:  BOONE, THOMAS CHARLES
;  TITLE OF INVENTION: MODIFIED PEPTIDES AS THERAPEUTIC AGENTS
;  FILE REFERENCE: A-527
;  CURRENT APPLICATION NUMBER: US/10/653,048
;  CURRENT FILING DATE:  2003-08-29
;  PRIOR APPLICATION NUMBER: US/09/428,082B
```

; PRIOR FILING DATE: 1999-10-22
; PRIOR APPLICATION NUMBER: 60/105,371
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 1133
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1030
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: EPO-MIMETIC PEPTIDE
US-10-653-048-1030

Query Match 93.4%; Score 57; DB 4; Length 20;
Best Local Similarity 57.1%; Pred. No. 0.012;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 YXCXXGPXTWXCXP 14
| | || || | |
Db 4 YACRMGPITWVCSP 17

RESULT 15
US-10-460-550-9
; Sequence 9, Application US/10460550
; Publication No. US20040092444A1
; GENERAL INFORMATION:
; APPLICANT: Digicaylioglu, Murat
; APPLICANT: Lipton, Stuart A.
; TITLE OF INVENTION: Neuroprotective Synergy of
; TITLE OF INVENTION: Erythropoietin and Insulin-Like Growth Factors
; FILE REFERENCE: 66821-210
; CURRENT APPLICATION NUMBER: US/10/460,550
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: US 60/388,058
; PRIOR FILING DATE: 2002-06-11
; PRIOR APPLICATION NUMBER: US 60/458,145
; PRIOR FILING DATE: 2003-03-26
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-10-460-550-9

Query Match 93.4%; Score 57; DB 4; Length 20;
Best Local Similarity 57.1%; Pred. No. 0.012;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 YXCXXGPXTWXCXP 14
| | || || | |
Db 4 YACRMGPITWVCSP 17

Search completed: March 31, 2006, 17:34:53
Job time : 43.1144 secs

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: March 31, 2006, 17:24:52 ; Search time 5.43284 Seconds
(without alignments)
78.446 Million cell updates/sec

Title: US-10-609-217-83
Perfect score: 61
Sequence: 1 YXCXXGPXTWXCXP 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 180808 seqs, 30441898 residues

Total number of hits satisfying chosen parameters: 180808

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA_New:*
1: /SIDS5/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
2: /SIDS5/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
3: /SIDS5/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
4: /SIDS5/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
5: /SIDS5/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
6: /SIDS5/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
7: /SIDS5/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
8: /SIDS5/ptodata/2/pubpaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	57	93.4	20	6	US-10-935-005B-4	Sequence 4, Appli
2	57	93.4	20	6	US-10-935-005B-25	Sequence 25, Appl
3	56	91.8	18	6	US-10-935-005B-13	Sequence 13, Appl
4	56	91.8	20	6	US-10-935-005B-2	Sequence 2, Appli
5	56	91.8	20	6	US-10-935-005B-7	Sequence 7, Appli
6	56	91.8	20	6	US-10-935-005B-14	Sequence 14, Appl
7	56	91.8	20	6	US-10-935-005B-23	Sequence 23, Appl
8	56	91.8	20	7	US-11-261-157-4	Sequence 4, Appli
9	56	91.8	20	7	US-11-007-772A-117	Sequence 117, App
10	56	91.8	23	6	US-10-935-005B-8	Sequence 8, Appli
11	56	91.8	247	6	US-10-935-005B-82	Sequence 82, Appl
12	56	91.8	247	6	US-10-935-005B-85	Sequence 85, Appl
13	56	91.8	247	6	US-10-935-005B-87	Sequence 87, Appl
14	56	91.8	247	6	US-10-935-005B-88	Sequence 88, Appl
15	56	91.8	249	6	US-10-935-005B-83	Sequence 83, Appl
16	56	91.8	249	6	US-10-935-005B-86	Sequence 86, Appl
17	56	91.8	249	6	US-10-935-005B-89	Sequence 89, Appl
18	56	91.8	251	6	US-10-935-005B-84	Sequence 84, Appl
19	55	90.2	14	6	US-10-935-005B-1	Sequence 1, Appli
20	55	90.2	20	6	US-10-935-005B-3	Sequence 3, Appli
21	55	90.2	20	6	US-10-935-005B-5	Sequence 5, Appli
22	55	90.2	20	6	US-10-935-005B-24	Sequence 24, Appl
23	55	90.2	22	6	US-10-935-005B-15	Sequence 15, Appl
24	55	90.2	22	6	US-10-935-005B-26	Sequence 26, Appl
25	50	82.0	13	6	US-10-935-005B-16	Sequence 16, Appl
26	50	82.0	13	6	US-10-935-005B-27	Sequence 27, Appl
27	50	82.0	18	6	US-10-935-005B-12	Sequence 12, Appl
28	49	80.3	18	6	US-10-935-005B-21	Sequence 21, Appl
29	49	80.3	20	6	US-10-935-005B-6	Sequence 6, Appli
30	49	80.3	20	6	US-10-935-005B-22	Sequence 22, Appl
31	46.5	76.2	19	6	US-10-935-005B-10	Sequence 10, Appl
32	46	75.4	19	6	US-10-935-005B-11	Sequence 11, Appl
33	45	73.8	213	7	US-11-096-568A-20805	Sequence 20805, A
34	43	70.5	10	6	US-10-935-005B-9	Sequence 9, Appli
35	43	70.5	11	6	US-10-935-005B-17	Sequence 17, Appl
36	43	70.5	11	6	US-10-935-005B-28	Sequence 28, Appl
37	43	70.5	12	6	US-10-935-005B-29	Sequence 29, Appl
38	43	70.5	20	7	US-11-261-157-1	Sequence 1, Appli
39	43	70.5	20	7	US-11-261-157-3	Sequence 3, Appli
40	43	70.5	20	7	US-11-261-157-11	Sequence 11, Appl
41	43	70.5	20	7	US-11-261-157-12	Sequence 12, Appl
42	43	70.5	20	7	US-11-261-157-13	Sequence 13, Appl
43	43	70.5	21	7	US-11-261-157-2	Sequence 2, Appli
44	39	63.9	57	7	US-11-096-568A-26223	Sequence 26223, A
45	38	62.3	285	7	US-11-072-512-2821	Sequence 2821, Ap

ALIGNMENTS

RESULT 1
US-10-935-005B-4
; Sequence 4, Application US/10935005B
; Publication No. US20060051844A1
; GENERAL INFORMATION:
; APPLICANT: HEAVNER, George A.; KNIGHT, David; GHRAYEB, John; SCALLON, Bernard;
; APPLICANT: NESSPOR, Thomas; HUANG, Chichang
; TITLE OF INVENTION: HUMAN EPO MIMETIC HINGE CORE MIMETIBODIES,
; FILE REFERENCE: CEN5039NP
; CURRENT APPLICATION NUMBER: US/10/935,005B
; CURRENT FILING DATE: 2004-09-03
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 4
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-935-005B-4

Query Match 93.4%; Score 57; DB 6; Length 20;
Best Local Similarity 57.1%; Pred. No. 0.0022;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
Qy          1 YXCXXGPXTWXCXP 14
              | |  || || | |
Db          4 YACRMGPITWVCSP 17
```

RESULT 2
US-10-935-005B-25
; Sequence 25, Application US/10935005B
; Publication No. US20060051844A1
; GENERAL INFORMATION:
; APPLICANT: HEAVNER, George A.; KNIGHT, David; GHRAYEB, John; SCALLON, Bernard;
; APPLICANT: NESSPOR, Thomas; HUANG, Chichang
; TITLE OF INVENTION: HUMAN EPO MIMETIC HINGE CORE MIMETIBODIES,
; FILE REFERENCE: CEN5039NP
; CURRENT APPLICATION NUMBER: US/10/935,005B
; CURRENT FILING DATE: 2004-09-03
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 25
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:

; OTHER INFORMATION: synthetic peptide
US-10-935-005B-25

Query Match 93.4%; Score 57; DB 6; Length 20;
Best Local Similarity 57.1%; Pred. No. 0.0022;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 YXCXXGPXTWXCXP 14
Db 4 YACRMGPITWVCSP 17

RESULT 3
US-10-935-005B-13
; Sequence 13, Application US/10935005B
; Publication No. US20060051844A1
; GENERAL INFORMATION:
; APPLICANT: HEAVNER, George A.; KNIGHT, David; GHRAYEB, John; SCALLON, Bernard;
; APPLICANT: NESSPOR, Thomas; HUANG, Chichang
; TITLE OF INVENTION: HUMAN EPO MIMETIC HINGE CORE MIMETIBODIES,
; FILE REFERENCE: CEN5039NP
; CURRENT APPLICATION NUMBER: US/10/935,005B
; CURRENT FILING DATE: 2004-09-03
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 13
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-935-005B-13

Query Match 91.8%; Score 56; DB 6; Length 18;
Best Local Similarity 57.1%; Pred. No. 0.003;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 YXCXXGPXTWXCXP 14
Db 4 YSCHFGPLTWVCKP 17

RESULT 4
US-10-935-005B-2
; Sequence 2, Application US/10935005B
; Publication No. US20060051844A1
; GENERAL INFORMATION:
; APPLICANT: HEAVNER, George A.; KNIGHT, David; GHRAYEB, John; SCALLON, Bernard;
; APPLICANT: NESSPOR, Thomas; HUANG, Chichang
; TITLE OF INVENTION: HUMAN EPO MIMETIC HINGE CORE MIMETIBODIES,
; FILE REFERENCE: CEN5039NP
; CURRENT APPLICATION NUMBER: US/10/935,005B
; CURRENT FILING DATE: 2004-09-03
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-935-005B-2

Query Match 91.8%; Score 56; DB 6; Length 20;
Best Local Similarity 57.1%; Pred. No. 0.0032;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 YXCXXGPXTWXCXP 14
Db 4 YSCHFGPLTWVCKP 17

RESULT 5
US-10-935-005B-7
; Sequence 7, Application US/10935005B
; Publication No. US20060051844A1
; GENERAL INFORMATION:
; APPLICANT: HEAVNER, George A.; KNIGHT, David; GHRAYEB, John; SCALLON, Bernard;
; APPLICANT: NESSPOR, Thomas; HUANG, Chichang
; TITLE OF INVENTION: HUMAN EPO MIMETIC HINGE CORE MIMETIBODIES,
; FILE REFERENCE: CEN5039NP
; CURRENT APPLICATION NUMBER: US/10/935,005B
; CURRENT FILING DATE: 2004-09-03
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 7
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-935-005B-7

Query Match 91.8%; Score 56; DB 6; Length 20;
Best Local Similarity 57.1%; Pred. No. 0.0032;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 YXCXXGPXTWXCXP 14
Db 4 YSCHFGPLTWVCKP 17

RESULT 6
US-10-935-005B-14
; Sequence 14, Application US/10935005B
; Publication No. US20060051844A1
; GENERAL INFORMATION:
; APPLICANT: HEAVNER, George A.; KNIGHT, David; GHRAYEB, John; SCALLON, Bernard;
; APPLICANT: NESSPOR, Thomas; HUANG, Chichang
; TITLE OF INVENTION: HUMAN EPO MIMETIC HINGE CORE MIMETIBODIES,
; FILE REFERENCE: CEN5039NP
; CURRENT APPLICATION NUMBER: US/10/935,005B
; CURRENT FILING DATE: 2004-09-03
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 14
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-935-005B-14

Query Match 91.8%; Score 56; DB 6; Length 20;
Best Local Similarity 57.1%; Pred. No. 0.0032;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 YXCXXGPXTWXCXP 14
Db 4 YACHMGPMTWVCQP 17

RESULT 7
US-10-935-005B-23
; Sequence 23, Application US/10935005B
; Publication No. US20060051844A1
; GENERAL INFORMATION:
; APPLICANT: HEAVNER, George A.; KNIGHT, David; GHRAYEB, John; SCALLON, Bernard;
; APPLICANT: NESSPOR, Thomas; HUANG, Chichang
; TITLE OF INVENTION: HUMAN EPO MIMETIC HINGE CORE MIMETIBODIES,
; FILE REFERENCE: CEN5039NP
; CURRENT APPLICATION NUMBER: US/10/935,005B
; CURRENT FILING DATE: 2004-09-03
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 23

```
;   LENGTH: 20
;   TYPE: PRT
;   ORGANISM: Artificial
;   FEATURE:
;   OTHER INFORMATION: synthetic peptide
US-10-935-005B-23

  Query Match             91.8%;  Score 56;  DB 6;  Length 20;
  Best Local Similarity   57.1%;  Pred. No. 0.0032;
  Matches    8;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 YXCXXGPXTWXCXP 14
              | |  || || | |
Db          4 YSCHFGPLTWVCKP 17


RESULT 8
US-11-261-157-4
; Sequence 4, Application US/11261157
; Publication No. US20060040858A1
; GENERAL INFORMATION:
;  APPLICANT: Holmes, Christopher P.
;  APPLICANT:  Yin, Qun
;  APPLICANT:  Lalande, Guy
;  APPLICANT:  Schatz, Peter J.
;  APPLICANT:  Tumelty, David
;  APPLICANT:  Palani, Balu
;  APPLICANT:  Zemede, Genet
;  TITLE OF INVENTION: NOVEL PEPTIDES THAT BIND TO THE ERYTHROPOIETIN RECEPTOR
;  FILE REFERENCE: 04279/100M615-US2
;  CURRENT APPLICATION NUMBER: US/11/261,157
;  CURRENT FILING DATE:  2005-10-27
;  PRIOR APPLICATION NUMBER: 60/469,993
;  PRIOR FILING DATE: 2003-05-12
;  PRIOR APPLICATION NUMBER: 60/470,244
;  PRIOR FILING DATE: 2003-05-12
;  PRIOR APPLICATION NUMBER: 10/844,968
;  PRIOR FILING DATE: 2004-05-12
;  NUMBER OF SEQ ID NOS: 13
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 4
;   LENGTH: 20
;   TYPE: PRT
;   ORGANISM: artificial
;   FEATURE:
;   OTHER INFORMATION: synthetic peptide
;   FEATURE:
;   NAME/KEY: MISC_FEATURE
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: biotinylated
US-11-261-157-4

  Query Match             91.8%;  Score 56;  DB 7;  Length 20;
  Best Local Similarity   57.1%;  Pred. No. 0.0032;
  Matches    8;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 YXCXXGPXTWXCXP 14
              | |  || || | |
Db          4 YACHMGPITWVCQP 17


RESULT 9
US-11-007-772A-117
; Sequence 117, Application US/11007772A
; Publication No. US20060063699A1
; GENERAL INFORMATION:
;  APPLICANT: Larsen, Bjarne Due
;  TITLE OF INVENTION: Pharmacologically Active Peptide Conjugates Having a Reduced
;  TITLE OF INVENTION:  Tendency Towards Enzymatic Hydrolysis.
;  FILE REFERENCE: 50412/008004
;  CURRENT APPLICATION NUMBER: US/11/007,772A
;  CURRENT FILING DATE:  2004-12-07
;  PRIOR APPLICATION NUMBER: 09/341,590
;  PRIOR FILING DATE: 1999-07-12
;  PRIOR APPLICATION NUMBER: PCT/DK99/00118
;  PRIOR FILING DATE: 1999-03-09
;  PRIOR APPLICATION NUMBER: DK 0317/98
;  PRIOR FILING DATE: 1998-03-09
;  NUMBER OF SEQ ID NOS: 134
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 117
;   LENGTH: 20
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic
US-11-007-772A-117

  Query Match             91.8%;  Score 56;  DB 7;  Length 20;
  Best Local Similarity   57.1%;  Pred. No. 0.0032;
  Matches    8;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 YXCXXGPXTWXCXP 14
              | |  || || | |
Db          4 YSCHFGPLTWVCKP 17


RESULT 10
US-10-935-005B-8
; Sequence 8, Application US/10935005B
; Publication No. US20060051844A1
; GENERAL INFORMATION:
;  APPLICANT: HEAVNER, George A.; KNIGHT, David; GHRAYEB, John; SCALLON, Bernard;
;  APPLICANT:  NESSPOR, Thomas; HUANG, Chichang
;  TITLE OF INVENTION: HUMAN EPO MIMETIC HINGE CORE MIMETIBODIES,
;  FILE REFERENCE: CEN5039NP
;  CURRENT APPLICATION NUMBER: US/10/935,005B
;  CURRENT FILING DATE:  2004-09-03
;  NUMBER OF SEQ ID NOS: 89
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 8
;   LENGTH: 23
;   TYPE: PRT
;   ORGANISM: Artificial
;   FEATURE:
;   OTHER INFORMATION: synthetic peptide
US-10-935-005B-8

  Query Match             91.8%;  Score 56;  DB 6;  Length 23;
  Best Local Similarity   57.1%;  Pred. No. 0.0035;
  Matches    8;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 YXCXXGPXTWXCXP 14
              | |  || || | |
Db          4 YSCHFGPLTWVCKP 17


RESULT 11
US-10-935-005B-82
; Sequence 82, Application US/10935005B
; Publication No. US20060051844A1
; GENERAL INFORMATION:
;  APPLICANT: HEAVNER, George A.; KNIGHT, David; GHRAYEB, John; SCALLON, Bernard;
;  APPLICANT:  NESSPOR, Thomas; HUANG, Chichang
;  TITLE OF INVENTION: HUMAN EPO MIMETIC HINGE CORE MIMETIBODIES,
;  FILE REFERENCE: CEN5039NP
;  CURRENT APPLICATION NUMBER: US/10/935,005B
;  CURRENT FILING DATE:  2004-09-03
;  NUMBER OF SEQ ID NOS: 89
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 82
;   LENGTH: 247
;   TYPE: PRT
;   ORGANISM: Artificial
;   FEATURE:
;   OTHER INFORMATION: synthetic peptide
```

US-10-935-005B-82

Query Match 91.8%; Score 56; DB 6; Length 247;
Best Local Similarity 57.1%; Pred. No. 0.016;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 YXCXXGPXTWXCXP 14
| | || || | |
Db 7 YSCHFGPLTWVCKP 20

RESULT 12
US-10-935-005B-85
; Sequence 85, Application US/10935005B
; Publication No. US20060051844A1
; GENERAL INFORMATION:
; APPLICANT: HEAVNER, George A.; KNIGHT, David; GHRAYEB, John; SCALLON, Bernard;
; APPLICANT: NESSPOR, Thomas; HUANG, Chichang
; TITLE OF INVENTION: HUMAN EPO MIMETIC HINGE CORE MIMETIBODIES,
; FILE REFERENCE: CEN5039NP
; CURRENT APPLICATION NUMBER: US/10/935,005B
; CURRENT FILING DATE: 2004-09-03
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 85
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-935-005B-85

Query Match 91.8%; Score 56; DB 6; Length 247;
Best Local Similarity 57.1%; Pred. No. 0.016;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 YXCXXGPXTWXCXP 14
| | || || | |
Db 7 YSCHFGPLTWVCKP 20

RESULT 13
US-10-935-005B-87
; Sequence 87, Application US/10935005B
; Publication No. US20060051844A1
; GENERAL INFORMATION:
; APPLICANT: HEAVNER, George A.; KNIGHT, David; GHRAYEB, John; SCALLON, Bernard;
; APPLICANT: NESSPOR, Thomas; HUANG, Chichang
; TITLE OF INVENTION: HUMAN EPO MIMETIC HINGE CORE MIMETIBODIES,
; FILE REFERENCE: CEN5039NP
; CURRENT APPLICATION NUMBER: US/10/935,005B
; CURRENT FILING DATE: 2004-09-03
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 87
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-935-005B-87

Query Match 91.8%; Score 56; DB 6; Length 247;
Best Local Similarity 57.1%; Pred. No. 0.016;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 YXCXXGPXTWXCXP 14
| | || || | |
Db 7 YSCHFGPLTWVCKP 20

RESULT 14
US-10-935-005B-88
; Sequence 88, Application US/10935005B
; Publication No. US20060051844A1
; GENERAL INFORMATION:
; APPLICANT: HEAVNER, George A.; KNIGHT, David; GHRAYEB, John; SCALLON, Bernard;
; APPLICANT: NESSPOR, Thomas; HUANG, Chichang
; TITLE OF INVENTION: HUMAN EPO MIMETIC HINGE CORE MIMETIBODIES,
; FILE REFERENCE: CEN5039NP
; CURRENT APPLICATION NUMBER: US/10/935,005B
; CURRENT FILING DATE: 2004-09-03
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 88
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-935-005B-88

Query Match 91.8%; Score 56; DB 6; Length 247;
Best Local Similarity 57.1%; Pred. No. 0.016;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 YXCXXGPXTWXCXP 14
| | || || | |
Db 7 YSCHFGPLTWVCKP 20

RESULT 15
US-10-935-005B-83
; Sequence 83, Application US/10935005B
; Publication No. US20060051844A1
; GENERAL INFORMATION:
; APPLICANT: HEAVNER, George A.; KNIGHT, David; GHRAYEB, John; SCALLON, Bernard;
; APPLICANT: NESSPOR, Thomas; HUANG, Chichang
; TITLE OF INVENTION: HUMAN EPO MIMETIC HINGE CORE MIMETIBODIES,
; FILE REFERENCE: CEN5039NP
; CURRENT APPLICATION NUMBER: US/10/935,005B
; CURRENT FILING DATE: 2004-09-03
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 83
; LENGTH: 249
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-935-005B-83

Query Match 91.8%; Score 56; DB 6; Length 249;
Best Local Similarity 57.1%; Pred. No. 0.016;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 YXCXXGPXTWXCXP 14
| | || || | |
Db 7 YSCHFGPLTWVCKP 20

Search completed: March 31, 2006, 17:36:17
Job time : 5.43284 secs

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: March 31, 2006, 16:22:51 ; Search time 17.4129 Seconds
(without alignments)
154.717 Million cell updates/sec

Title: US-10-609-217-84
Perfect score: 122
Sequence: 1 YXCXXGPXTWXCXPYXCXXGPXTWXCXP 28

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	56	45.9	369	2	I46531	surfactant protein
2	56	45.9	370	1	LNRBB	pulmonary surfacta
3	53	43.4	415	2	PC4407	envelope protein -
4	48	39.3	287	2	A31876	myogenin - rat
5	48	39.3	363	2	A29072	pulmonary surfacta
6	48	39.3	381	1	LNHUB	pulmonary surfacta
7	47.5	38.9	82	2	C69013	hypothetical prote
8	46	37.7	138	2	PC1197	genome polyprotein
9	45.5	37.3	341	1	PYVZCB	spheroidin precurs
10	45	36.9	19	1	EWSMAN	ancovenin - Strept
11	45	36.9	135	2	A48776	polyprotein (E2/NS
12	43.5	35.7	1599	2	T16210	hypothetical prote
13	43	35.2	422	2	D86446	hypothetical prote
14	43	35.2	428	2	JQ1864	hypothetical 47.0K
15	43	35.2	1487	2	G96827	protein F20B17.10
16	42.5	34.8	116	2	C70656	hypothetical prote
17	42.5	34.8	368	2	A45831	MHC class I histoc
18	42.5	34.8	439	2	S51939	chitinase (EC 3.2.
19	42.5	34.8	4660	2	T42737	gp330 protein prec
20	42	34.4	136	2	S24090	envelope protein -
21	42	34.4	138	2	S24068	envelope protein -
22	42	34.4	138	2	S24084	envelope protein -
23	42	34.4	138	2	S24096	envelope protein -
24	42	34.4	296	2	T35345	chitinase - Strept
25	42	34.4	550	2	JH0711	genome polyprotein
26	42	34.4	723	2	T38780	hypothetical prote
27	42	34.4	1531	1	DVHUAR	multidrug resistan
28	41.5	34.0	283	2	S34851	hypothetical 31.9K
29	41.5	34.0	4543	1	A53102	alpha-2-macroglobu
30	41	33.6	134	2	S24089	envelope protein -
31	41	33.6	135	2	S24086	envelope protein -
32	41	33.6	136	2	S24091	envelope protein -
33	41	33.6	137	2	S24088	envelope protein -
34	41	33.6	138	2	S24102	envelope protein -
35	41	33.6	138	2	S24073	envelope protein -
36	41	33.6	138	2	S24069	envelope protein -
37	41	33.6	138	2	S24107	envelope protein -
38	41	33.6	138	2	S24087	envelope protein -
39	41	33.6	138	2	S24070	envelope protein -
40	41	33.6	138	2	S24078	envelope protein -
41	41	33.6	138	2	S24105	envelope protein -
42	41	33.6	138	2	PC1182	genome polyprotein
43	41	33.6	138	2	S24097	envelope protein -
44	41	33.6	854	1	QRHYLD	LDL receptor precu
45	41	33.6	860	1	QRHULD	LDL receptor precu

ALIGNMENTS

RESULT 1
I46531
surfactant protein B - rabbit
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 14-Feb-1997 #sequence_revision 14-Feb-1997 #text_change 09-Jul-2004
C;Accession: I46531
R;Margana, R.K.; Boggaram, V.
Am. J. Physiol. 268, L481-L490, 1995
A;Title: Transcription and mRNA stability regulate developmental and hormonal expression
A;Reference number: I46531; MUID:95208794; PMID:7900830
A;Accession: I46531
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-369 <MAR>
A;Cross-references: UNIPROT:P15285; UNIPARC:UPI000016C51A; EMBL:U17106; NID:g642487; PID:
C;Genetics:
A;Gene: SP-B
C;Superfamily: pulmonary surfactant protein B; saposin repeat homology
F;61-153/Domain: saposin repeat homology <SAP1>

Query Match 45.9%; Score 56; DB 2; Length 369;
Best Local Similarity 36.4%; Pred. No. 0.35;
Matches 8; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 3 CXXGPXTWXCXPYXCXXGPXTW 24
Db 18 CGPGTAVWATSPLACAQGPEFW 39

RESULT 2
LNRBB
pulmonary surfactant protein B precursor - rabbit
N;Alternate names: pulmonary surfactant-associated protein-B
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 09-Jul-2004
C;Accession: A32421
R;Xu, J.; Richardson, C.; Ford, C.; Spencer, T.; Li-juan, Y.; Mackie, G.; Hammond, G.; P
Biochem. Biophys. Res. Commun. 160, 325-332, 1989
A;Title: Isolation and characterization of the cDNA for pulmonary surfactant-associated
A;Reference number: A32421; MUID:89228033; PMID:2469419
A;Accession: A32421
A;Molecule type: mRNA
A;Residues: 1-370 <XUJ>
A;Cross-references: UNIPROT:P15285; UNIPARC:UPI000016C5F3; GB:M24901; NID:g165707; PIDN:
A;Note: the authors translated the codon CCG for residue 184 as Arg and CAG for residue
C;Comment: Pulmonary surfactant is a complex of phospholipids and proteins that lowers t
C;Superfamily: pulmonary surfactant protein B; saposin repeat homology
C;Keywords: alveolar proteinosis; gaseous exchange; glycoprotein; lipoprotein; lung; pul
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-184/Domain: propeptide #status predicted <PRO>
F;62-154/Domain: saposin repeat homology <SAP1>
F;184-271/Domain: saposin repeat homology <SAP2>

```
F;185-263/Product: pulmonary surfactant protein B, 9K form #status predicted <SP9>
F;185-240/Product: pulmonary surfactant protein B, 6K form #status predicted <SP6>
F;280-365/Domain: saposin repeat homology <SAP3>
F;300/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             45.9%;  Score 56;  DB 1;  Length 370;
  Best Local Similarity   36.4%;  Pred. No. 0.35;
  Matches    8;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy          3 CXXGPXTWXCXPYXCXXGPXTW 24
              |  |   |   |  |  ||  |
Db         19 CGPGTAVWATSPLACAQGPEFW 40


RESULT 3
PC4407
envelope protein - hepatitis C virus (fragment)
C;Species: hepatitis C virus
C;Date: 10-Nov-1997 #sequence_revision 23-Jan-1998 #text_change 09-Jul-2004
C;Accession: PC4407
R;Li, G.; Yao, J.; Peng, W.
Chinese J. Virol. 13, 24-32, 1997
A;Title: Sequence of genomic region of hepatitis C virus envelope proteins from a Guangd
A;Reference number: PC4407
A;Accession: PC4407
A;Molecule type: genomic RNA
A;Residues: 1-415 <LIA>
A;Cross-references: UNIPROT:Q7LZY4; UNIPARC:UPI0000178545
A;Note: the authors translated the codon ATA for residues 93 and 249 as Met
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: envelope protein

  Query Match             43.4%;  Score 53;  DB 2;  Length 415;
  Best Local Similarity   36.4%;  Pred. No. 1.1;
  Matches    8;  Conservative    1;  Mismatches   13;  Indels    0;  Gaps    0;

Qy          7 PXTWXCXPYXCXXGPXTWXCXP 28
              |  |   |   |   | :| | |
Db        319 PYCWHYAPRPCGIVPASWVCGP 340


RESULT 4
A31876
myogenin - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 08-Jun-1989 #sequence_revision 08-Jun-1989 #text_change 09-Jul-2004
C;Accession: A31876
R;Wright, W.E.; Sassoon, D.A.; Lin, V.K.
Cell 56, 607-617, 1989
A;Title: Myogenin, a factor regulating myogenesis, has a domain homologous to MyoD.
A;Reference number: A31876; MUID:89136007; PMID:2537150
A;Accession: A31876
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-287 <WRI>
A;Cross-references: UNIPROT:P20428; UNIPARC:UPI000012FBA6; GB:M24393; NID:g205604; PIDN:
C;Keywords: DNA binding

  Query Match             39.3%;  Score 48;  DB 2;  Length 287;
  Best Local Similarity   33.3%;  Pred. No. 4.2;
  Matches    8;  Conservative    1;  Mismatches   15;  Indels    0;  Gaps    0;

Qy          3 CXXGPXTWXCXPYXCXXGPXTWXC 26
              |   | :|   |      ||  | |
Db        232 CAWEPLSWCQTPPLLQQGPFKWGC 255


RESULT 5
A29072
pulmonary surfactant protein SP 18 precursor - dog (fragment)
C;Species: Canis lupus familiaris (dog)
C;Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 09-Jul-2004
C;Accession: B29072; A29072
R;Hawgood, S.; Benson, B.J.; Schilling, J.; Damm, D.; Clements, J.A.; White, R.T.
Proc. Natl. Acad. Sci. U.S.A. 84, 66-70, 1987
A;Title: Nucleotide and amino acid sequences of pulmonary surfactant protein SP 18 and ev
A;Reference number: A29072; MUID:87092398; PMID:3467361
A;Accession: B29072
A;Molecule type: mRNA
A;Residues: 1-363 <HAW>
A;Cross-references: UNIPROT:P17129; UNIPARC:UPI00001327F3; GB:M15170; NID:g164077; PIDN:A
A;Accession: A29072
A;Molecule type: protein
A;Residues: 182-210 <HA2>
A;Cross-references: UNIPARC:UPI0000177937
C;Superfamily: pulmonary surfactant protein B; saposin repeat homology
F;1-14/Domain: signal sequence #status predicted <SIG>
F;15-180/Domain: propeptide #status predicted <PRP>
F;54-146/Domain: saposin repeat homology <SAP1>
F;180-267/Domain: saposin repeat homology <SAP2>
F;181-363/Product: pulmonary surfactant protein SP 18 #status experimental <MAT>
F;273-358/Domain: saposin repeat homology <SAP3>

  Query Match             39.3%;  Score 48;  DB 2;  Length 363;
  Best Local Similarity   31.8%;  Pred. No. 5;
  Matches    7;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy          3 CXXGPXTWXCXPYXCXXGPXTW 24
              |  |   |       |  ||  |
Db         11 CGLGAADWSAPSLACARGPAFW 32


RESULT 6
LNHUB
pulmonary surfactant protein B precursor [validated] - human
N;Alternate names: pulmonary surfactant proteolipid SP-B; pulmonary surfactant-associatec
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 09-Jul-2004
C;Accession: A31361; A28461; A27794; A27592; JU0162; S21134
R;Pilot-Matias, T.J.; Kister, S.E.; Fox, J.L.; Kropp, K.; Glasser, S.W.; Whitsett, J.A.
DNA 8, 75-86, 1989
A;Title: Structure and organization of the gene encoding human pulmonary surfactant prote
A;Reference number: A31361; MUID:89170128; PMID:2924687
A;Accession: A31361
A;Molecule type: DNA
A;Residues: 1-381 <PIL>
A;Cross-references: UNIPROT:P07988; UNIPARC:UPI00001423D4; GB:M24461
A;Note: the codon given for residue 131 (ATT) is inconsistent with the authors' translati
A;Note: this protein is encoded by a single gene
R;Jacobs, K.A.; Phelps, D.S.; Steinbrink, R.; Fisch, J.; Kriz, R.; Mitsock, L.; Dougherty
J. Biol. Chem. 262, 9808-9811, 1987
A;Title: Isolation of a cDNA clone encoding a high molecular weight precursor to a 6-kDa
A;Reference number: A28461; MUID:87250653; PMID:3597440
A;Accession: A28461
A;Molecule type: mRNA
A;Residues: 1-227,'A',229-381 <JAC>
A;Cross-references: UNIPARC:UPI000000162D; GB:J02761; NID:g190673; PIDN:AAA60212.1; PID:g
A;Note: part of this sequence, including the amino end of the mature protein, was confirm
R;Glasser, S.W.; Korfhagen, T.R.; Weaver, T.; Pilot-Matias, T.; Fox, J.L.; Whitsett, J.A
Proc. Natl. Acad. Sci. U.S.A. 84, 4007-4011, 1987
A;Title: cDNA and deduced amino acid sequence of human pulmonary surfactant-associated pr
A;Reference number: A27794; MUID:87231940; PMID:3035561
A;Accession: A27794
A;Molecule type: mRNA
A;Residues: 'EFR',99-317,'L',319-381 <GLA>
A;Cross-references: UNIPARC:UPI000014237D; GB:M16764; NID:g338410; PIDN:AAA88099.1; PID:g
A;Note: 131-Ile was also found
A;Note: part of this sequence, including the amino end of the mature protein, was confirm
R;Revak, S.D.; Merritt, T.A.; Degryse, E.; Stefani, L.; Courtney, M.; Hallman, M.; Cochra
J. Clin. Invest. 81, 826-833, 1988
A;Title: Use of human surfactant low molecular weight apoproteins in the reconstitution c
A;Reference number: A27592; MUID:88139786; PMID:3343343
A;Accession: A27592
A;Molecule type: mRNA
A;Residues: 139-177,'V',179-227,'A',228-381 <REV>
A;Cross-references: UNIPARC:UPI00001741A7; GB:M19097
```

```
A;Note: part of this sequence, including the amino end of the mature protein, was confir
A;Note: the mature protein (SP 18) consists of two identical disulfide-bonded 9K polypep
R;Mizumoto, M.; Adachi, H.
Sapporo Igaku Zasshi 56, 731-742, 1987
A;Title: Primary structure of a hydrophobic 6kDa apoprotein (SP6) of human pulmonary sur
A;Reference number: JU0162
A;Accession: JU0162
A;Molecule type: protein
A;Residues: 201-207,'X',209-210,'X',212-227,'A',229-234,'X',236-245,'X',247,'L',249-253,
A;Cross-references: UNIPARC:UPI00001741A8
R;Johansson, J.; Joernvall, H.; Curstedt, T.
FEBS Lett. 301, 165-167, 1992
A;Title: Human surfactant polypeptide SP-B. Disulfide bridges, C-terminal end, and pepti
A;Reference number: S21134; MUID:92233937; PMID:1568474
A;Accession: S21134
A;Status: preliminary
A;Molecule type: protein
A;Residues: 201-227,'I',229-279 <JOH>
A;Cross-references: UNIPARC:UPI00001741A9
A;Note: 228-Ala was also found
C;Comment: Pulmonary surfactant is a complex of phospholipids and proteins that lowers t
C;Genetics:
A;Gene: GDB:SFTPB; SFTP3; SP-B
A;Cross-references: GDB:120374; OMIM:178640
A;Map position: 2p12-2p11.2
A;Introns: 23/1; 65/3; 89/3; 131/3; 194/3; 224/3; 286/1; 334/3; 361/3
C;Superfamily: pulmonary surfactant protein B; saposin repeat homology
C;Keywords: alveolar proteinosis; gaseous exchange; glycoprotein; lipoprotein; lung; pul
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-200/Domain: propeptide #status predicted <PRO>
F;61-153/Domain: saposin repeat homology <SAP1>
F;200-287/Domain: saposin repeat homology <SAP2>
F;201-279/Product: pulmonary surfactant protein B, 9K form #status predicted <SP9>
F;201-256/Product: pulmonary surfactant protein B, 6K form #status experimental <SP6>
F;291-376/Domain: saposin repeat homology <SAP3>
F;69-143,72-137,100-112,299-366,302-360,325-335/Disulfide bonds: #status predicted
F;129,311/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;208-277,211-271,235-246/Disulfide bonds: #status experimental
F;248/Disulfide bonds: interchain #status experimental

  Query Match             39.3%;  Score 48;  DB 1;  Length 381;
  Best Local Similarity   31.8%;  Pred. No. 5.2;
  Matches    7;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy          3 CXXGPXTWXCXPYXCXXGPXTW 24
              |  |   |      |  ||  |
Db         18 CGPGTAAWTTSSLACAQGPEFW 39


RESULT 7
C69013
hypothetical protein MTH110 - Methanobacterium thermoautotrophicum (strain Delta H)
C;Species: Methanobacterium thermoautotrophicum
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C;Accession: C69013
R;Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.
ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A;Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct
A;Reference number: A69000; MUID:98037514; PMID:9371463
A;Accession: C69013
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-82 <MTH>
A;Cross-references: UNIPROT:O26213; UNIPARC:UPI0000062AA2; GB:AE000801; GB:AE000666; NID
A;Experimental source: strain Delta H
C;Genetics:
A;Gene: MTH110
A;Start codon: GTG

  Query Match             38.9%;  Score 47.5;  DB 2;  Length 82;
  Best Local Similarity   35.0%;  Pred. No. 1.8;
  Matches    7;  Conservative    1;  Mismatches    9;  Indels    3;  Gaps    1;

Qy         10 WXCXPYXCXXGPXT---WXC 26
              | | |: |    |     | ||
Db         25 WVCAPFGCGGSPLCPIFWDC 44


RESULT 8
PC1197
genome polyprotein - hepatitis C virus (strain RS1-1) (fragment)
C;Species: hepatitis C virus
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C;Accession: PC1197
R;Kato, N.; Ootsuyama, Y.; Ohkoshi, S.; Nakazawa, T.; Sekiya, H.; Hijikata, M.; Shimotoh
Biochem. Biophys. Res. Commun. 189, 119-127, 1992
A;Title: Characterization of hypervariable regions in the putative envelope protein of h
A;Reference number: PC1182; MUID:93080545; PMID:1333186
A;Accession: PC1197
A;Status: nucleic acid sequence not shown
A;Molecule type: genomic RNA
A;Residues: 1-138 <KAT>
A;Cross-references: UNIPROT:Q81363; UNIPARC:UPI00000F9BF3; GB:D12957; NID:g285854; PIDN:
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: polyprotein

  Query Match             37.7%;  Score 46;  DB 2;  Length 138;
  Best Local Similarity   34.8%;  Pred. No. 4.5;
  Matches    8;  Conservative    1;  Mismatches   14;  Indels    0;  Gaps    0;

Qy          6 GPXTWXCXPYXCXXGPXTWXCXP 28
              ||  |   |  |   | :  | |
Db        115 GPYCWHYAPRPCGIVPASQVCGP 137


RESULT 9
PYVZCB
spheroidin precursor - Choristoneura biennis poxvirus
C;Species: Choristoneura biennis poxvirus
C;Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 09-Jul-2004
C;Accession: A34743
R;Yuen, L.; Dionne, J.; Arif, B.; Richardson, C.
Virology 175, 427-433, 1990
A;Title: Identification and sequencing of the spheroidin gene of Choristoneura biennis e
A;Reference number: A34743; MUID:90223988; PMID:2327073
A;Accession: A34743
A;Molecule type: DNA
A;Residues: 1-341 <YUE>
A;Cross-references: UNIPROT:P23061; UNIPARC:UPI0000135E39; EMBL:M34140; NID:g323273; PID
C;Comment: This protein appears to be essential for viral replication.
C;Superfamily: spheroidin
C;Keywords: glycoprotein; replication
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-341/Product: spheroidin #status predicted <SPD>
F;176,196/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             37.3%;  Score 45.5;  DB 1;  Length 341;
  Best Local Similarity   29.6%;  Pred. No. 11;
  Matches    8;  Conservative    2;  Mismatches   14;  Indels    3;  Gaps    1;

Qy          1 YXCXXGPXTWXCXPYXCXXG---PXTW 24
              | |  |   |   |:    |   | :|
Db        112 YLCAAGASDWSIRPFGDKSGMDLPGSW 138


RESULT 10
EWSMAN
ancovenin - Streptomyces sp. (strain A647P-2)
C;Species: Streptomyces sp.
C;Date: 12-May-1994 #sequence_revision 19-May-1994 #text_change 09-Jul-2004
C;Accession: A61284
R;Wakamiya, T.; Ueki, Y.; Shiba, T.; Kido, Y.; Motoki, Y.
Tetrahedron Lett. 26, 665-668, 1985
A;Title: The structure of ancovenin, a new peptide inhibitor of angiotensin I converting
```

```
A;Reference number: A61284
A;Accession: A61284
A;Molecule type: protein
A;Residues: 1-19 <WAK>
A;Cross-references: UNIPROT:P38655; UNIPARC:UPI0000052CC3
C;Superfamily: cinnamycin precursor
C;Keywords: antibiotic; lanthionine
F;1-18/Cross-link: (2S,3S,6R)-3-methyl-lanthionine (Cys-Thr) #status experimental
F;4-14/Cross-link: sn-(2S,6R)-lanthionine (Ser-Cys) #status experimental
F;5-11/Cross-link: (2S,3S,6R)-3-methyl-lanthionine (Cys-Thr) #status experimental
F;6/Modified site: dehydroalanine (Ser) #status experimental

  Query Match             36.9%;  Score 45;  DB 1;  Length 19;
  Best Local Similarity   60.0%;  Pred. No. 1.3;
  Matches    6;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          3 CXXGPXTWXC 12
              |  || || |
Db          5 CSFGPLTWSC 14


RESULT 11
A48776
polyprotein (E2/NS1 region, HVR1, HVR2) - hepatitis C virus (fragment)
C;Species: hepatitis C virus
C;Date: 07-Apr-1994 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: A48776
R;Higashi, Y.; Kakumu, S.; Yoshioka, K.; Wakita, T.; Mizokami, M.; Ohba, K.; Ito, Y.; Is
Virology 197, 659-668, 1993
A;Title: Dynamics of genome change in the E2/NS1 region of hepatitis C virus in vivo.
A;Reference number: A48776; MUID:94069940; PMID:8249288
A;Accession: A48776
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: nucleic acid
A;Residues: 1-135 <HIG>
A;Cross-references: UNIPROT:Q9PXU1; UNIPARC:UPI00000EE867
A;Experimental source: subtype II, patient FJ
A;Note: sequence extracted from NCBI backbone (NCBIP:140212)
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: polyprotein

  Query Match             36.9%;  Score 45;  DB 2;  Length 135;
  Best Local Similarity   31.8%;  Pred. No. 6.2;
  Matches    7;  Conservative    2;  Mismatches   13;  Indels    0;  Gaps    0;

Qy          7 PXTWXCXPYXCXXGPXTWXCXP 28
              |  |   |: |   | :  | |
Db        113 PYCWHYAPFPCGNVPASQVCGP 134


RESULT 12
T16210
hypothetical protein F30H5.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C;Accession: T16210
R;Pauley, A.; Stellyes, L.
submitted to the EMBL Data Library, June 1995
A;Description: The sequence of C. elegans cosmid F30H5.
A;Reference number: Z18478
A;Accession: T16210
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1599 <PAU>
A;Cross-references: UNIPROT:Q09983; UNIPARC:UPI000007E415; EMBL:U29096; NID:g861390; PID
A;Experimental source: strain Bristol N2
C;Genetics:
A;Gene: CESP:F30H5.3
A;Introns: 12/1; 59/2; 85/3; 124/3; 217/2; 534/3; 560/1; 1549/1

  Query Match             35.7%;  Score 43.5;  DB 2;  Length 1599;
  Best Local Similarity   33.3%;  Pred. No. 75;
  Matches    8;  Conservative    0;  Mismatches   13;  Indels    3;  Gaps    1;

Qy          1 YXCXXGPXTWXCXP---YXCXXGP 21
              | |      | | |   | |   |
Db        651 YECYFDGYVWGCCPTKAYTCTLSP 674


RESULT 13
D86446
hypothetical protein F3C3.3 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: D86446
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: D86446
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-422 <STO>
A;Cross-references: UNIPROT:Q9FVR4; UNIPARC:UPI000000C253; GB:AE005172; NID:g10801376; PI
C;Genetics:
A;Map position: 1
C;Superfamily: Arabidopsis thaliana hypothetical protein F18O22.180

  Query Match             35.2%;  Score 43;  DB 2;  Length 422;
  Best Local Similarity   30.8%;  Pred. No. 30;
  Matches    8;  Conservative    2;  Mismatches   16;  Indels    0;  Gaps    0;

Qy          1 YXCXXGPXTWXCXPYXCXXGPXTWXC 26
              : |   | |  |    |  |  :| |
Db        374 WKCLKCPDTECCRSSCCCSGCFSWLC 399


RESULT 14
JQ1864
hypothetical 47.0K protein - bovine adenovirus 3
C;Species: Mastadenovirus bos3 (bovine adenovirus 3)
C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 09-Jul-2004
C;Accession: JQ1864
R;Mittal, S.K.; Prevec, L.; Babiuk, L.A.; Graham, F.L.
J. Gen. Virol. 73, 3295-3300, 1992
A;Title: Sequence analysis of bovine adenovirus type 3 early region 3 and fibre protein g
A;Reference number: PQ0499; MUID:93107871; PMID:1469367
A;Accession: JQ1864
A;Molecule type: DNA
A;Residues: 1-428 <MIT>
A;Cross-references: UNIPROT:O71105; UNIPARC:UPI0000179E82; DDBJ:D12928
A;Experimental source: strain WBR-1
A;Note: the authors described carbohydrate binding site for residue 67

  Query Match             35.2%;  Score 43;  DB 2;  Length 428;
  Best Local Similarity   50.0%;  Pred. No. 31;
  Matches    6;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          6 GPXTWXCXPYXC 17
              |  || | |: |
Db        186 GNVTWFCPPFMC 197


RESULT 15
G96827
protein F20B17.10 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: G96827
```

R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: G96827
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1487 <STO>
A;Cross-references: UNIPROT:Q9MA08; UNIPARC:UPI000009DAA5; GB:AE005173; NID:g7715604; PI
C;Genetics:
A;Gene: F20B17.10
A;Map position: 1

Query Match 35.2%; Score 43; DB 2; Length 1487;
Best Local Similarity 29.2%; Pred. No. 83;
Matches 7; Conservative 1; Mismatches 16; Indels 0; Gaps 0;

```
Qy          3 CXXGPXTWXCXPYXCXXGPXTWXC 26
              |  |     |    |   | :| |
Db       1018 CEEGKGLSSCGELTCVNVPGSWRC 1041
```

Search completed: March 31, 2006, 16:37:16
Job time : 18.4129 secs

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: March 31, 2006, 16:09:36 ; Search time 104.896 Seconds
(without alignments)
188.328 Million cell updates/sec

Title: US-10-609-217-84
Perfect score: 122
Sequence: 1 YXCXXGPXTWXCXPYXCXXGPXTWXCXP 28

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	56	45.9	370	1	PSPB_RABIT	P15285 oryctolagus
2	53	43.4	352	2	Q4IMN3_GIBZE	Q4imn3 gibberella
3	53	43.4	415	2	Q7LZY4_9HEPC	Q7lzy4 hepatitis c
4	50	41.0	671	2	Q4IDR8_GIBZE	Q4idr8 gibberella
5	48.5	39.8	544	2	Q4SD11_TETNG	Q4sd11 tetraodon n
6	48.5	39.8	891	2	Q7T2X3_CHICK	Q7t2x3 gallus gall
7	48	39.3	247	2	Q4R8P3_MACFA	Q4r8p3 macaca fasc
8	48	39.3	287	1	MYOG_RAT	P20428 rattus norv
9	48	39.3	363	1	PSPB_CANFA	P17129 canis famil
10	48	39.3	381	1	PSPB_HUMAN	P07988 homo sapien
11	48	39.3	1322	2	Q4SDN8_TETNG	Q4sdn8 tetraodon n
12	47.5	38.9	82	2	O26213_METTH	O26213 methanobact
13	47.5	38.9	389	2	Q8C5R5_MOUSE	Q8c5r5 mus musculu
14	47	38.5	180	2	Q4I355_GIBZE	Q4i355 gibberella
15	46	37.7	137	2	Q9YK68_9HEPC	Q9yk68 hepatitis c
16	46	37.7	138	2	Q81363_9HEPC	Q81363 hepatitis c
17	46	37.7	414	2	Q4SAV9_TETNG	Q4sav9 tetraodon n
18	45.5	37.3	210	2	Q7Q2J1_ANOGA	Q7q2j1 anopheles g
19	45.5	37.3	341	1	SPIN_CBEPV	P23061 choristoneu
20	45	36.9	19	1	DURC_STRGP	P36503 streptomyce
21	45	36.9	19	1	LANC_STRS6	P38655 streptomyce
22	45	36.9	135	2	Q9PXU1_9HEPC	Q9pxu1 hepatitis c
23	45	36.9	151	2	Q9YK30_9HEPC	Q9yk30 hepatitis c
24	45	36.9	151	2	Q9YK46_9HEPC	Q9yk46 hepatitis c
25	45	36.9	176	2	Q5MM97_9HEPC	Q5mm97 hepatitis c
26	45	36.9	229	2	Q6SJQ2_MOUSE	Q6sjq2 mus musculu
27	45	36.9	426	2	Q5XMX9_9HEPC	Q5xmx9 hepatitis c
28	45	36.9	504	2	Q7QWR4_GIALA	Q7qwr4 giardia lam
29	45	36.9	729	2	Q7T3M4_BRARE	Q7t3m4 brachydanio
30	45	36.9	729	2	Q4V9K5_BRARE	Q4v9k5 brachydanio
31	44.5	36.5	339	2	Q9YJU4_CFEPV	Q9yju4 choristoneu
32	44	36.1	46	2	Q81408_9HEPC	Q81408 hepatitis c
33	44	36.1	141	2	Q4T8M8_TETNG	Q4t8m8 tetraodon n
34	44	36.1	186	2	Q9IXW7_9HEPC	Q9ixw7 hepatitis c
35	44	36.1	186	2	Q9IYL7_9HEPC	Q9iyl7 hepatitis c
36	44	36.1	202	2	Q6ZV28_HUMAN	Q6zv28 homo sapien
37	44	36.1	215	2	Q4WFS9_ASPFU	Q4wfs9 aspergillus
38	44	36.1	245	2	Q6ZNJ8_HUMAN	Q6znj8 homo sapien
39	44	36.1	335	2	Q91HQ1_9HEPC	Q91hq1 hepatitis c
40	44	36.1	418	2	Q8TDT0_HUMAN	Q8tdt0 homo sapien
41	43.5	35.7	284	2	Q4W0V0_HUMAN	Q4w0v0 homo sapien
42	43.5	35.7	1599	2	Q616G7_CAEBR	Q616g7 caenorhabdi
43	43.5	35.7	1599	2	Q09983_CAEEL	Q09983 caenorhabdi
44	43	35.2	61	2	O70227_RAT	O70227 rattus norv
45	43	35.2	191	2	Q8BBK9_9HEPC	Q8bbk9 hepatitis c

ALIGNMENTS

```
RESULT 1
PSPB_RABIT
ID   PSPB_RABIT     STANDARD;      PRT;   370 AA.
AC   P15285; P79333;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   10-MAY-2005 (Rel. 47, Last annotation update)
DE   Pulmonary surfactant-associated protein B precursor (SP-B) (6 kDa
DE   protein) (Pulmonary surfactant-associated proteolipid SPL(Phe)).
GN   Name=SFTPB; Synonyms=SFTP3;
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Lagomorpha; Leporidae;
OC   Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [MRNA].
RC   TISSUE=Lung;
RX   MEDLINE=89228033; PubMed=2469419;
RA   Xu J., Richardson C., Ford C., Spencer T., Li-Juan Y., Mackie G.,
RA   Hammond G., Possmayer F.;
RT   "Isolation and characterization of the cDNA for pulmonary surfactant-
RT   associated protein-B (SP-B) in the rabbit.";
RL   Biochem. Biophys. Res. Commun. 160:325-332(1989).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=New Zealand white;
RX   MEDLINE=95208794; PubMed=7900830;
RA   Margana R.K., Boggaram V.;
RT   "Transcription and mRNA stability regulate developmental and hormonal
RT   expression of rabbit surfactant protein B gene.";
RL   Am. J. Physiol. 268:L481-L490(1995).
RN   [3]
RP   NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RC   TISSUE=Liver;
RX   MEDLINE=96198312; PubMed=8928820;
RA   Margana R.K., Boggaram V.;
RT   "Rabbit surfactant protein B gene: structure and functional
RT   characterization of the promoter.";
RL   Am. J. Physiol. 270:L601-L612(1996).
RN   [4]
RP   NUCLEOTIDE SEQUENCE OF 1-34.
RX   MEDLINE=96096536; PubMed=8522191; DOI=10.1016/0378-1119(95)00576-R;
RA   Luzi P., Anceschi M., Strayer D.S.;
RT   "The upstream region of the SP-B gene: intrinsic promoter activity and
RT   glucocorticoid responsiveness related to a new DNA-binding protein.";
RL   Gene 165:285-290(1995).
CC   -!- FUNCTION: Pulmonary surfactant-associated proteins promote
CC       alveolar stability by lowering the surface tension at the air-
CC       liquid interface in the peripheral air spaces. SP-B increases the
CC       collapse pressure of palmitic acid to nearly 70 millinewtons per
CC       meter.
CC   -!- SUBUNIT: Homodimer; disulfide-linked.
CC   -!- SUBCELLULAR LOCATION: Secreted; extracellular.
```

```
CC   -!- MISCELLANEOUS: Pulmonary surfactant consists of 90% lipid and 10%
CC       protein. There are 4 surfactant-associated protein: 2 collagenous,
CC       carbohydrate-binding glycoproteins (SP-A and SP-D) and 2 small
CC       hydrophobic proteins (SP-B and SP-C).
CC   -!- SIMILARITY: Contains 1 saposin A-type domain.
CC   -!- SIMILARITY: Contains 3 saposin B-type domains.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; M24901; AAA31466.1; -; mRNA.
DR   EMBL; U17106; AAA67934.1; -; mRNA.
DR   EMBL; U40853; AAB48076.1; -; Genomic_DNA.
DR   EMBL; S80649; AAD14335.1; -; Genomic_DNA.
DR   PIR; A32421; LNRBB.
DR   PIR; I46531; I46531.
DR   HSSP; P07988; 1DFW.
DR   InterPro; IPR009007; Pept_Aspartc_cat.
DR   InterPro; IPR003119; SapA.
DR   InterPro; IPR007856; SapB_1.
DR   InterPro; IPR008138; SapB_2.
DR   InterPro; IPR008373; Saposin.
DR   InterPro; IPR008139; SaposinB.
DR   InterPro; IPR008137; Surfactant_B.
DR   Pfam; PF02199; SapA; 1.
DR   Pfam; PF05184; SapB_1; 1.
DR   Pfam; PF03489; SapB_2; 3.
DR   PRINTS; PR01797; SAPOSIN.
DR   ProDom; PD008002; Surfactant_B; 1.
DR   SMART; SM00162; SAPA; 1.
DR   SMART; SM00741; SapB; 3.
DR   PROSITE; PS51110; SAP_A; 1.
DR   PROSITE; PS50015; SAP_B; 3.
KW   Gaseous exchange; Glycoprotein; Repeat; Surface film.
FT   PROPEP          1    184
FT   CHAIN         185    263       Pulmonary surfactant-associated protein
FT                                   B.
FT   PROPEP        264    370
FT   DOMAIN         26     66       Saposin A-type.
FT   DOMAIN         66    148       Saposin B-type 1.
FT   DOMAIN        188    263       Saposin B-type 2.
FT   DOMAIN        284    359       Saposin B-type 3.
FT   CARBOHYD      300    300       N-linked (GlcNAc. . .) (Potential).
FT   DISULFID      192    261       By similarity.
FT   DISULFID      195    255       By similarity.
FT   DISULFID      219    230       By similarity.
FT   DISULFID      232    232       Interchain (By similarity).
FT   CONFLICT       15     15       Missing (in Ref. 2).
FT   CONFLICT      129    129       Q -> L (in Ref. 3).
FT   CONFLICT      184    184       R -> P (in Ref. 1).
FT   CONFLICT      232    232       C -> R (in Ref. 2).
FT   CONFLICT      289    289       R -> P (in Ref. 3).
FT   CONFLICT      329    355       ELHTPQLLSLLSRGWDARAICQALGAC -> AAHAPAAEPA
FT                                   VQGLGCPRNLPGPEGRV (in Ref. 1).
SQ   SEQUENCE   370 AA;  40610 MW;  423047A69B12DCB5 CRC64;

  Query Match             45.9%;  Score 56;  DB 1;  Length 370;
  Best Local Similarity   36.4%;  Pred. No. 0.78;
  Matches    8;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy          3 CXXGPXTWXCXPYXCXXGPXTW 24
              | |   |   |  |  || |
Db         19 CGPGTAVWATSPLACAQGPEFW 40


RESULT 2
Q4IMN3_GIBZE
ID   Q4IMN3_GIBZE PRELIMINARY;      PRT;   352 AA.
AC   Q4IMN3;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Hypothetical protein.
GN   ORFNames=FG01525.1;
OS   Gibberella zeae PH-1.
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC   Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
OX   NCBI_TaxID=229533;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=PH-1;
RA   Birren B., Nusbaum C., Abouelleil A., Allen N., Anderson S.,
RA   Arachchi H.M., Barna N., Bastien V., Bloom T., Boguslavkiy L.,
RA   Boukhgalter B., Butler J., Calvo S.E., Camarata J., Chang J.,
RA   Choepel Y., Collymore A., Cook A., Cooke P., Corum B., DeArellano K.,
RA   Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
RA   Erickson J., Faro S., Ferreira P., FitzGerald M., Gage D., Galagan J.,
RA   Gardyna S., Gnerre S., Graham L., Grand-Pierre N., Hafez N.,
RA   Hagopian D., Hagos B., Hall J., Horton L., Hulme W., Iliev I.,
RA   Jaffe D., Johnson R., Jones C., Kamal M., Kamat A., Karatas A.,
RA   Kells C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
RA   Ma L.-J., Mabbitt R., MacLean C., Macdonald P., Major J., Manning J.,
RA   Matthews C., Mauceli E., McCarthy M., Meldrim J., Meneus L.,
RA   Mihova T., Mlenga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
RA   Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neil D.,
RA   Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,
RA   Rachupka A., Ramasamy U., Raymond C., Retta R., Rise C., Rogov P.,
RA   Roman J., Schauer S., Schupback R., Seaman S., Severy P., Smirnov S.,
RA   Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,
RA   Talamas J., Tesfaye S., Theodore J., Topham K., Travers M.,
RA   Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,
RA   Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
RA   Lander E.;
RT   "Fusarium graminearum genome sequence.";
RL   Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; AACM01000077; EAA68151.1; -; Genomic_DNA.
KW   Hypothetical protein.
SQ   SEQUENCE   352 AA;  38308 MW;  670BA49FC645A7F8 CRC64;

  Query Match             43.4%;  Score 53;  DB 2;  Length 352;
  Best Local Similarity   53.8%;  Pred. No. 2.2;
  Matches    7;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          3 CXXGPXTWXCXPY 15
              |   | || | ||
Db        184 CTSNPSTWRCYPY 196


RESULT 3
Q7LZY4_9HEPC
ID   Q7LZY4_9HEPC PRELIMINARY;      PRT;   415 AA.
AC   Q7LZY4;
DT   01-MAR-2004 (TrEMBLrel. 26, Created)
DT   01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Envelope protein (Fragment).
OS   Hepatitis C virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Li G., Yao J., Peng W.;
RT   "Sequence of genomic region of hepatitis C virus envelope proteins
RT   from a Guangdong patient.";
RL   Submitted (NOV-1997) to the PIR data bank.
DR   PIR; PC4407; PC4407.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
```

```
DR   InterPro; IPR002521; HCV_core.
DR   InterPro; IPR002519; HCV_env.
DR   InterPro; IPR002531; HCV_NS1.
DR   Pfam; PF01542; HCV_core; 1.
DR   Pfam; PF01539; HCV_env; 1.
DR   Pfam; PF01560; HCV_NS1; 1.
KW   Envelope protein; Transmembrane.
FT   NON_TER       1      1
FT   NON_TER     415    415
SQ   SEQUENCE   415 AA;  44999 MW;  2CD9FEE53B33AB92 CRC64;

  Query Match            43.4%;  Score 53;  DB 2;  Length 415;
  Best Local Similarity  36.4%;  Pred. No. 2.6;
  Matches    8;  Conservative    1;  Mismatches   13;  Indels    0;  Gaps    0;

Qy          7 PXTWXCXPYXCXXGPXTWXCXP 28
              |  |   |  |    | :| | |
Db        319 PYCWHYAPRPCGIVPASWVCGP 340


RESULT 4
Q4IDR8_GIBZE
ID   Q4IDR8_GIBZE PRELIMINARY;      PRT;   671 AA.
AC   Q4IDR8;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Hypothetical protein.
GN   ORFNames=FG04640.1;
OS   Gibberella zeae PH-1.
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC   Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
OX   NCBI_TaxID=229533;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=PH-1;
RA   Birren B., Nusbaum C., Abouelleil A., Allen N., Anderson S.,
RA   Arachchi H.M., Barna N., Bastien V., Bloom T., Boguslavkiy L.,
RA   Boukhgalter B., Butler J., Calvo S.E., Camarata J., Chang J.,
RA   Choepel Y., Collymore A., Cook A., Cooke P., Corum B., DeArellano K.,
RA   Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
RA   Erickson J., Faro S., Ferreira P., FitzGerald M., Gage D., Galagan J.,
RA   Gardyna S., Gnerre S., Graham L., Grand-Pierre N., Hafez N.,
RA   Hagopian D., Hagos B., Hall J., Horton L., Hulme W., Iliev I.,
RA   Jaffe D., Johnson R., Jones C., Kamal M., Kamat A., Karatas A.,
RA   Kells C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
RA   Ma L.-J., Mabbitt R., MacLean C., Macdonald P., Major J., Manning J.,
RA   Matthews C., Mauceli E., McCarthy M., Meldrim J., Meneus L.,
RA   Mihova T., Mlenga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
RA   Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neil D.,
RA   Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,
RA   Rachupka A., Ramasamy U., Raymond C., Retta R., Rise C., Rogov P.,
RA   Roman J., Schauer S., Schupback R., Seaman S., Severy P., Smirnov S.,
RA   Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,
RA   Talamas J., Tesfaye S., Theodore J., Topham K., Travers M.,
RA   Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,
RA   Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
RA   Lander E.;
RT   "Fusarium graminearum genome sequence.";
RL   Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; AACM01000194; EAA72557.1; -; Genomic_DNA.
KW   Hypothetical protein.
SQ   SEQUENCE   671 AA;  76165 MW;  AA44789EEEFB2030 CRC64;

  Query Match            41.0%;  Score 50;  DB 2;  Length 671;
  Best Local Similarity  31.8%;  Pred. No. 12;
  Matches    7;  Conservative    1;  Mismatches   14;  Indels    0;  Gaps    0;

Qy          7 PXTWXCXPYXCXXGPXTWXCXP 28
              |  | | |: |      |   |


Db        563 PSVWNCEPFTCSHTFDVWPIRP 584


RESULT 5
Q4SD11_TETNG
ID   Q4SD11_TETNG PRELIMINARY;      PRT;   544 AA.
AC   Q4SD11;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Chromosome 14 SCAF14645, whole genome shotgun sequence.
DE   (Fragment).
GN   ORFNames=GSTENG00020243001;
OS   Tetraodon nigroviridis (Green puffer).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC   Tetradontoidea; Tetraodontidae; Tetraodon.
OX   NCBI_TaxID=99883;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA   Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA   Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA   Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA   Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA   Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA   Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA   Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA   Kellis M., Volff JN., Guigo R., Zody M.C., Mesirov J.,
RA   Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA   Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA   Wincker P., Lander E.S., Weissenbach J., Roest Crollius H.;
RT   "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT   the early vertebrate proto-karyotype.";
RL   Nature 431:946-957(2004).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RG   Genoscope; Whitehead Institute Centre for Genome Research;
RL   Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
CC   -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
DR   EMBL; CAAE01014645; CAG01471.1; -; Genomic_DNA.
DR   InterPro; IPR001356; Homeobox.
DR   Pfam; PF00046; Homeobox; 2.
DR   PRINTS; PR00024; HOMEOBOX.
DR   ProDom; PD000010; Homeobox; 2.
DR   SMART; SM00389; HOX; 2.
DR   PROSITE; PS00027; HOMEOBOX_1; 2.
DR   PROSITE; PS50071; HOMEOBOX_2; 2.
KW   DNA-binding; Homeobox; Nuclear protein.
FT   NON_TER     544    544
SQ   SEQUENCE   544 AA;  58130 MW;  EBAE01DAF17CB05E CRC64;

  Query Match            39.8%;  Score 48.5;  DB 2;  Length 544;
  Best Local Similarity  34.5%;  Pred. No. 17;
  Matches   10;  Conservative    1;  Mismatches   15;  Indels    3;  Gaps    1;

Qy          3 CXXGPXTW---XCXPYXCXXGPXTWXCXP 28
              |  || ||       : |  ||    | |
Db        137 CGTGPRTWRRSSAPAWPCRCGPAPRCCPP 165


RESULT 6
Q7T2X3_CHICK
ID   Q7T2X3_CHICK PRELIMINARY;      PRT;   891 AA.
AC   Q7T2X3;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Low-density lipoprotein receptor precursor.
```

```
GN   Name=LDLR;
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Schneider W.J.;
RT   "Molecular characterization of the chicken LDL receptor.";
RL   Submitted (SEP-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ515243; CAD56163.1; -; mRNA.
DR   HSSP; P01130; 1F5Y.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0005509; F:calcium ion binding; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR000742; EGF_2.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR002172; LDL_receptor_A.
DR   InterPro; IPR000033; Ldl_receptor_rep.
DR   Pfam; PF00008; EGF; 2.
DR   Pfam; PF07645; EGF_CA; 1.
DR   Pfam; PF00057; Ldl_recept_a; 7.
DR   Pfam; PF00058; Ldl_recept_b; 4.
DR   PRINTS; PR00261; LDLRECEPTOR.
DR   SMART; SM00179; EGF_CA; 2.
DR   SMART; SM00192; LDLa; 7.
DR   SMART; SM00135; LY; 5.
DR   PROSITE; PS00010; ASX_HYDROXYL; 2.
DR   PROSITE; PS01186; EGF_2; 2.
DR   PROSITE; PS50026; EGF_3; 2.
DR   PROSITE; PS01187; EGF_CA; 2.
DR   PROSITE; PS01209; LDLRA_1; 6.
DR   PROSITE; PS50068; LDLRA_2; 7.
KW   Lipoprotein; Receptor; Signal.
FT   SIGNAL       1     15       Potential.
FT   CHAIN       16    891       low-density lipoprotein receptor.
SQ   SEQUENCE   891 AA;  93986 MW;  D483C146496E7BB6 CRC64;

  Query Match             39.8%;  Score 48.5;  DB 2;  Length 891;
  Best Local Similarity   34.8%;  Pred. No. 26;
  Matches    8;  Conservative    2;  Mismatches   12;  Indels    1;  Gaps    1;

Qy          7 PXTWXCXPY-XCXXGPXTWXCXP 28
              | :| |  :  |  |   | | |
Db         80 PLSWRCDGHRDCRHGADEWGCEP 102


RESULT 7
Q4R8P3_MACFA
ID   Q4R8P3_MACFA PRELIMINARY;      PRT;   247 AA.
AC   Q4R8P3;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Testis cDNA clone: QtsA-11907, similar to human SEC7 homolog
DE   (TIC),.
OS   Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
OC   Cercopithecidae; Cercopithecinae; Macaca.
OX   NCBI_TaxID=9541;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   International consortium for macaque cDNA sequencing, analysis;
RT   "DNA sequences of macaque genes expressed in brain or testis and its
RT   evolutionary implications.";
RL   Submitted (JUN-2005) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RA   Osada N., Hirata M., Tanuma R., Kusuda J., Hida M., Suzuki Y.,
RA   Sugano S., Gojobori T., Shen J.C.-K., Wu C.I., Hashimoto K.;
RT   "Substitution rate and structural divergence of 5'UTR evolution:
RT   Comparative analysis between human and cynomolgus monkey cDNAs.";
RL   Submitted (MAR-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB168407; BAE00528.1; -; mRNA.
SQ   SEQUENCE   247 AA;  27884 MW;  E626250D77421039 CRC64;

  Query Match             39.3%;  Score 48;  DB 2;  Length 247;
  Best Local Similarity   38.9%;  Pred. No. 9.6;
  Matches    7;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          9 TWXCXPYXCXXGPXTWXC 26
              || |    |   |  | ||
Db        212 TWSCWWPACTAPPMLWTC 229


RESULT 8
MYOG_RAT
ID   MYOG_RAT       STANDARD;      PRT;   287 AA.
AC   P20428;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   10-MAY-2005 (Rel. 47, Last annotation update)
DE   Myogenin.
GN   Name=Myog;
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC   Muroidea; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=89136007; PubMed=2537150;
RA   Wright W.E., Sassoon D.A., Lin V.K.;
RT   "Myogenin, a factor regulating myogenesis, has a domain homologous to
RT   MyoD.";
RL   Cell 56:607-617(1989).
CC   -!- FUNCTION: Involved in muscle differentiation (myogenic factor).
CC       Induces fibroblasts to differentiate into myoblasts. Probable
CC       sequence specific DNA-binding protein.
CC   -!- SUBUNIT: Efficient DNA binding requires dimerization with another
CC       bHLH protein.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC   ----------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ----------------------------------------------------------------------
DR   EMBL; M24393; AAA41662.1; -; mRNA.
DR   PIR; A31876; A31876.
DR   HSSP; P10085; 1MDY.
DR   SMR; P20428; 74-136.
DR   TRANSFAC; T01532; -.
DR   Ensembl; ENSRNOG00000030743; Rattus norvegicus.
DR   RGD; 620432; Myog.
DR   InterPro; IPR002546; Basic.
DR   InterPro; IPR001092; HLH_basic.
DR   Pfam; PF01586; Basic; 1.
DR   Pfam; PF00010; HLH; 1.
DR   SMART; SM00520; BASIC; 1.
DR   SMART; SM00353; HLH; 1.
DR   PROSITE; PS50888; HLH; 1.
KW   Developmental protein; Differentiation; DNA-binding; Myogenesis;
KW   Nuclear protein.
FT   DOMAIN       94    133       Helix-loop-helix motif.
FT   DNA_BIND     81     93       Basic motif.
SQ   SEQUENCE   287 AA;  32503 MW;  BE454E59B1464B40 CRC64;

  Query Match             39.3%;  Score 48;  DB 1;  Length 287;
  Best Local Similarity   33.3%;  Pred. No. 11;
```

```
  Matches    8;  Conservative    1;  Mismatches   15;  Indels    0;  Gaps    0;

Qy          3 CXXGPXTWXCXPYXCXXGPXTWXC 26
              |   | :|   |     ||  | ||
Db        232 CAWEPLSWCQTPPLLQQGPFKWGC 255


RESULT 9
PSPB_CANFA
ID   PSPB_CANFA     STANDARD;      PRT;   363 AA.
AC   P17129;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   10-MAY-2005 (Rel. 47, Last annotation update)
DE   Pulmonary surfactant-associated protein B precursor (SP-B) (6 kDa
DE   protein) (Pulmonary surfactant-associated proteolipid SPL(Phe))
DE   (Pulmonary surfactant protein 18) (SP 18) (Fragment).
GN   Name=SFTPB; Synonyms=SFTP3;
OS   Canis familiaris (Dog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
OC   Canis.
OX   NCBI_TaxID=9615;
RN   [1]
RP   NUCLEOTIDE SEQUENCE, AND PROTEIN SEQUENCE OF 182-211.
RC   TISSUE=Lung;
RX   MEDLINE=87092398; PubMed=3467361;
RA   Hawgood S., Benson B.J., Schilling J., Damm D., Clements J.A.,
RA   White R.T.;
RT   "Nucleotide and amino acid sequences of pulmonary surfactant protein
RT   SP 18 and evidence for cooperation between SP 18 and SP 28-36 in
RT   surfactant lipid adsorption.";
RL   Proc. Natl. Acad. Sci. U.S.A. 84:66-70(1987).
CC   -!- FUNCTION: Pulmonary surfactant-associated proteins promote
CC       alveolar stability by lowering the surface tension at the air-
CC       liquid interface in the peripheral air spaces. SP-B increases the
CC       collapse pressure of palmitic acid to nearly 70 millinewtons per
CC       meter.
CC   -!- SUBUNIT: Homodimer; disulfide-linked.
CC   -!- SUBCELLULAR LOCATION: Secreted; extracellular.
CC   -!- MISCELLANEOUS: Pulmonary surfactant consists of 90% lipid and 10%
CC       protein. There are 4 surfactant-associated protein: 2 collagenous,
CC       carbohydrate-binding glycoproteins (SP-A and SP-D) and 2 small
CC       hydrophobic proteins (SP-B and SP-C).
CC   -!- SIMILARITY: Contains 1 saposin A-type domain.
CC   -!- SIMILARITY: Contains 3 saposin B-type domains.
CC   ----------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ----------------------------------------------------------------------
DR   EMBL; M15170; AAA30893.1; -; mRNA.
DR   PIR; B29072; A29072.
DR   HSSP; P07988; 1DFW.
DR   Ensembl; ENSCAFG00000007658; Canis familiaris.
DR   InterPro; IPR009007; Pept_Aspartc_cat.
DR   InterPro; IPR003119; SapA.
DR   InterPro; IPR007856; SapB_1.
DR   InterPro; IPR008138; SapB_2.
DR   InterPro; IPR008373; Saposin.
DR   InterPro; IPR008139; SaposinB.
DR   InterPro; IPR008137; Surfactant_B.
DR   Pfam; PF02199; SapA; 1.
DR   Pfam; PF05184; SapB_1; 1.
DR   Pfam; PF03489; SapB_2; 3.
DR   PRINTS; PR01797; SAPOSIN.
DR   ProDom; PD008002; Surfactant_B; 1.
DR   SMART; SM00162; SAPA; 1.
DR   SMART; SM00741; SapB; 3.
DR   PROSITE; PS51110; SAP_A; 1.
DR   PROSITE; PS50015; SAP_B; 3.
KW   Direct protein sequencing; Gaseous exchange; Glycoprotein; Repeat;
KW   Surface film.
FT   PROPEP       <1    180
FT   CHAIN       181    259       Pulmonary surfactant-associated protein
FT                                 B.
FT   PROPEP      260    363
FT   DOMAIN       18     58       Saposin A-type.
FT   DOMAIN       58    140       Saposin B-type 1.
FT   DOMAIN      184    259       Saposin B-type 2.
FT   DOMAIN      277    352       Saposin B-type 3.
FT   CARBOHYD    293    293       N-linked (GlcNAc. . .) (Potential).
FT   DISULFID    188    257       By similarity.
FT   DISULFID    191    251       By similarity.
FT   DISULFID    215    226       By similarity.
FT   DISULFID    228    228       Interchain (By similarity).
FT   NON_TER       1      1
SQ   SEQUENCE   363 AA;  40180 MW;  F4DAD0E02D8B2719 CRC64;

  Query Match             39.3%;  Score 48;  DB 1;  Length 363;
  Best Local Similarity   31.8%;  Pred. No. 14;
  Matches    7;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy          3 CXXGPXTWXCXPYXCXXGPXTW 24
              |  |  |     |  ||  |
Db         11 CGLGAADWSAPSLACARGPAFW 32


RESULT 10
PSPB_HUMAN
ID   PSPB_HUMAN     STANDARD;      PRT;   381 AA.
AC   P07988; Q96R04;
DT   01-AUG-1988 (Rel. 08, Created)
DT   01-MAY-1992 (Rel. 22, Last sequence update)
DT   10-MAY-2005 (Rel. 47, Last annotation update)
DE   Pulmonary surfactant-associated protein B precursor (SP-B) (6 kDa
DE   protein) (Pulmonary surfactant-associated proteolipid SPL(Phe)) (18
DE   kDa pulmonary-surfactant protein).
GN   Name=SFTPB; Synonyms=SFTP3;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [MRNA], AND PROTEIN SEQUENCE OF 201-214.
RC   TISSUE=Lung;
RX   MEDLINE=87250653; PubMed=3597440;
RA   Jacobs K.A., Phelps D.S., Steinbrink R., Fisch J., Kriz R.,
RA   Mitsock L., Dougherty J.P., Taeusch H.W., Floros J.;
RT   "Isolation of a cDNA clone encoding a high molecular weight precursor
RT   to a 6-kDa pulmonary surfactant-associated protein.";
RL   J. Biol. Chem. 262:9808-9811(1987).
RN   [2]
RP   NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RX   MEDLINE=89170128; PubMed=2924687;
RA   Pilot-Matias T.J., Kister S.E., Fox J.L., Kropp K., Glasser S.W.,
RA   Whitsett J.A.;
RT   "Structure and organization of the gene encoding human pulmonary
RT   surfactant proteolipid SP-B.";
RL   DNA 8:75-86(1989).
RN   [3]
RP   NUCLEOTIDE SEQUENCE [GENOMIC DNA], AND VARIANTS ILE-131; PHE-176 AND
RP   HIS-272.
RA   Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,
RA   Nickerson D.A.;
RT   "SeattleSNPs. NHLBI HL66682 program for genomic applications, UW-
RT   FHCRC, Seattle, WA (URL: http://pga.gs.washington.edu).";
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC   TISSUE=Brain;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
```

```
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [5]
RP   NUCLEOTIDE SEQUENCE [MRNA] OF 99-381.
RX   MEDLINE=87231940; PubMed=3035561;
RA   Glasser S.W., Korfhagen T.R., Weaver T., Pilot-Matias T., Fox J.L.,
RA   Whitsett J.A.;
RT   "cDNA and deduced amino acid sequence of human pulmonary surfactant-
RT   associated proteolipid SPL(Phe).";
RL   Proc. Natl. Acad. Sci. U.S.A. 84:4007-4011(1987).
RN   [6]
RP   NUCLEOTIDE SEQUENCE OF 139-381.
RX   MEDLINE=88139786; PubMed=3343343;
RA   Revak S.D., Merritt T.A., Degryse E., Stefani L., Courtney M.,
RA   Hallman M., Cochrane C.G.;
RT   "Use of human surfactant low molecular weight apoproteins in the
RT   reconstitution of surfactant biologic activity.";
RL   J. Clin. Invest. 81:826-833(1988).
RN   [7]
RP   PROTEIN SEQUENCE OF 201-279, AND DISULFIDE BONDS.
RX   MEDLINE=92233937; PubMed=1568474; DOI=10.1016/0014-5793(92)81239-I;
RA   Johansson J., Joernvall H., Curstedt T.;
RT   "Human surfactant polypeptide SP-B. Disulfide bridges, C-terminal end,
RT   and peptide analysis of the airway form.";
RL   FEBS Lett. 301:165-167(1992).
RN   [8]
RP   STRUCTURE BY FTIR OF 201-225.
RX   MEDLINE=20256237; PubMed=10798379;
RX   DOI=10.1034/j.1399-3011.2000.00693.x;
RA   Gordon L.M., Lee K.Y., Lipp M.M., Zasadzinski J.A., Walther F.J.,
RA   Sherman M.A., Waring A.J.;
RT   "Conformational mapping of the N-terminal segment of surfactant
RT   protein B in lipid using 13C-enhanced Fourier transform infrared
RT   spectroscopy.";
RL   J. Pept. Res. 55:330-347(2000).
RN   [9]
RP   VARIANT ILE-131.
RX   MEDLINE=20525299; PubMed=11076040;
RA   Lin Z., Pearson C., Chinchilli V., Pietschmann S.M., Luo J., Pison U.,
RA   Floros J.;
RT   "Polymorphisms of human SP-A, SP-B, and SP-D genes: association of SP-
RT   B Thr131Ile with ARDS.";
RL   Clin. Genet. 58:181-191(2000).
CC   -!- FUNCTION: Pulmonary surfactant-associated proteins promote
CC       alveolar stability by lowering the surface tension at the air-
CC       liquid interface in the peripheral air spaces. SP-B increases the
CC       collapse pressure of palmitic acid to nearly 70 millinewtons per
CC       meter.
CC   -!- SUBUNIT: Homodimer; disulfide-linked.
CC   -!- SUBCELLULAR LOCATION: Secreted; extracellular.
CC   -!- POLYMORPHISM: There seems to be an association between the Ile-131
CC       polymorphism and acute respiratory distress syndrome (ARDS).
CC   -!- DISEASE: Defects in SFTPB are a cause of congenital pulmonary
CC       alveolar proteinosis (PAP) [MIM:265120]. PAP is an autosomal
CC       recessive fatal respiratory disease.
CC   -!- MISCELLANEOUS: Pulmonary surfactant consists of 90% lipid and 10%
CC       protein. There are 4 surfactant-associated protein: 2 collagenous,
CC       carbohydrate-binding glycoproteins (SP-A and SP-D) and 2 small
CC       hydrophobic proteins (SP-B and SP-C).
CC   -!- SIMILARITY: Contains 1 saposin A-type domain.
CC   -!- SIMILARITY: Contains 3 saposin B-type domains.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; J02761; AAA60212.1; -; mRNA.
DR   EMBL; M24461; AAB59541.1; -; Genomic_DNA.
DR   EMBL; AF400074; AAK77913.1; -; Genomic_DNA.
DR   EMBL; BC032785; AAH32785.1; -; mRNA.
DR   EMBL; M16764; AAA88099.1; ALT_INIT; mRNA.
DR   EMBL; M19097; AAA36628.1; -; mRNA.
DR   PIR; A31361; LNHUB.
DR   PDB; 1DFW; IR; A=201-225.
DR   Ensembl; ENSG00000168878; Homo sapiens.
DR   HGNC; HGNC:10801; SFTPB.
DR   H-InvDB; HIX0002227; -.
DR   MIM; 178640; -.
DR   MIM; 265120; -.
DR   GO; GO:0009887; P:organogenesis; TAS.
DR   GO; GO:0007585; P:respiratory gaseous exchange; TAS.
DR   InterPro; IPR009007; Pept_Aspartc_cat.
DR   InterPro; IPR003119; SapA.
DR   InterPro; IPR007856; SapB_1.
DR   InterPro; IPR008138; SapB_2.
DR   InterPro; IPR008373; Saposin.
DR   InterPro; IPR008139; SaposinB.
DR   InterPro; IPR008137; Surfactant_B.
DR   Pfam; PF02199; SapA; 1.
DR   Pfam; PF05184; SapB_1; 1.
DR   Pfam; PF03489; SapB_2; 3.
DR   PRINTS; PR01797; SAPOSIN.
DR   ProDom; PD008002; Surfactant_B; 1.
DR   SMART; SM00162; SAPA; 1.
DR   SMART; SM00741; SapB; 3.
DR   PROSITE; PS51110; SAP_A; 1.
DR   PROSITE; PS50015; SAP_B; 3.
KW   3D-structure; Direct protein sequencing; Gaseous exchange;
KW   Glycoprotein; Polymorphism; Repeat; Surface film.
FT   PROPEP      1    200
FT   CHAIN     201    279       Pulmonary surfactant-associated protein
FT                              B.
FT   PROPEP    280    381
FT   DOMAIN     25     65       Saposin A-type.
FT   DOMAIN     65    147       Saposin B-type 1.
FT   DOMAIN    204    279       Saposin B-type 2.
FT   DOMAIN    295    370       Saposin B-type 3.
FT   CARBOHYD  129    129       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD  311    311       N-linked (GlcNAc. . .) (Potential).
FT   DISULFID  208    277
FT   DISULFID  211    271
FT   DISULFID  235    246
FT   DISULFID  248    248       Interchain.
FT   VARIANT   131    131       T -> I (in dbSNP:1130866).
FT                              /FTId=VAR_006948.
FT   VARIANT   176    176       L -> F (in dbSNP:3024801).
FT                              /FTId=VAR_013099.
FT   VARIANT   228    228       A -> I.
FT                              /FTId=VAR_006950.
FT   VARIANT   228    228       A -> R.
FT                              /FTId=VAR_006949.
FT   VARIANT   272    272       R -> H (in dbSNP:3024809).
FT                              /FTId=VAR_013100.
FT   CONFLICT  178    178       L -> V (in Ref. 6).
FT   CONFLICT  318    318       P -> L (in Ref. 5).
FT   HELIX     208    221
SQ   SEQUENCE   381 AA;  42117 MW;  9FD7F66678A35153 CRC64;
```

```
  Query Match             39.3%;  Score 48;  DB 1;  Length 381;
  Best Local Similarity   31.8%;  Pred. No. 14;
  Matches    7;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy          3 CXXGPXTWXCXPYXCXXGPXTW 24
              |  |   |      |  ||  |
Db         18 CGPGTAAWTTSSLACAQGPEFW 39


RESULT 11
Q4SDN8_TETNG
ID   Q4SDN8_TETNG PRELIMINARY;      PRT;  1322 AA.
AC   Q4SDN8;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Chromosome 10 SCAF14634, whole genome shotgun sequence.
GN   ORFNames=GSTENG00019945001;
OS   Tetraodon nigroviridis (Green puffer).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC   Tetradontoidea; Tetraodontidae; Tetraodon.
OX   NCBI_TaxID=99883;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA   Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA   Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA   Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA   Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA   Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA   Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA   Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA   Kellis M., Volff JN., Guigo R., Zody M.C., Mesirov J.,
RA   Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA   Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA   Wincker P., Lander E.S., Weissenbach J., Roest Crollius H.;
RT   "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT   the early vertebrate proto-karyotype.";
RL   Nature 431:946-957(2004).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RG   Genoscope; Whitehead Institute Centre for Genome Research;
RL   Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; CAAE01014634; CAG01244.1; -; Genomic_DNA.
SQ   SEQUENCE   1322 AA;  143378 MW;  49F57DB1F4849F1D CRC64;

  Query Match             39.3%;  Score 48;  DB 2;  Length 1322;
  Best Local Similarity   34.8%;  Pred. No. 46;
  Matches    8;  Conservative    2;  Mismatches   13;  Indels    0;  Gaps    0;

Qy          6 GPXTWXCXPYXCXXGPXTWXCXP 28
              || :|   |     || :|   |
Db       1131 GPGSWVLGPGPWVLGPGSWALGP 1153


RESULT 12
O26213_METTH
ID   O26213_METTH PRELIMINARY;      PRT;    82 AA.
AC   O26213;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Hypothetical protein MTH110.
GN   OrderedLocusNames=MTH110;
OS   Methanobacterium thermoautotrophicum.
OC   Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC   Methanobacteriaceae; Methanothermobacter.
OX   NCBI_TaxID=187420;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Delta H;
RX   MEDLINE=98037514; PubMed=9371463;
RA   Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
RA   Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA   Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
RA   Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA   Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA   McDougall S., Shimer G., Goyal A., Pietrovski S., Church G.M.,
RA   Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RT   "Complete genome sequence of Methanobacterium thermoautotrophicum
RT   deltaH: functional analysis and comparative genomics.";
RL   J. Bacteriol. 179:7135-7155(1997).
DR   EMBL; AE000801; AAB84616.1; -; Genomic_DNA.
DR   PIR; C69013; C69013.
KW   Complete proteome.
SQ   SEQUENCE   82 AA;  9325 MW;  A94449C8938F6198 CRC64;

  Query Match             38.9%;  Score 47.5;  DB 2;  Length 82;
  Best Local Similarity   35.0%;  Pred. No. 4.1;
  Matches    7;  Conservative    1;  Mismatches    9;  Indels    3;  Gaps    1;

Qy         10 WXCXPYXCXXGPXT---WXC 26
              | | |: |   |      | |
Db         25 WVCAPPGCGGSPLCPIFWDC 44


RESULT 13
Q8C5R5_MOUSE
ID   Q8C5R5_MOUSE PRELIMINARY;      PRT;   389 AA.
AC   Q8C5R5;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Mus musculus adult male testis cDNA, RIKEN full-length enriched
DE   library, clone:4933440E22 product:hypothetical protein, full insert
DE   sequence.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC   Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=C57BL/6J; TISSUE=Testis;
RX   MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA   Carninci P., Hayashizaki Y.;
RT   "High-efficiency full-length cDNA cloning.";
RL   Meth. Enzymol. 303:19-44(1999).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=C57BL/6J; TISSUE=Testis;
RX   MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
```

```
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=C57BL/6J; TISSUE=Testis;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
RN   [4]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=C57BL/6J; TISSUE=Testis;
RX   MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA   Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA   Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT   "Normalization and subtraction of cap-trapper-selected cDNAs to
RT   prepare full-length cDNA libraries for rapid discovery of new genes.";
RL   Genome Res. 10:1617-1630(2000).
RN   [5]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=C57BL/6J; TISSUE=Testis;
RX   MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA   Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA   Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,
RA   Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA   Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA   Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA   Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA   Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT   "RIKEN integrated sequence analysis (RISA) system-384-format
RT   sequencing pipeline with 384 multicapillary sequencer.";
RL   Genome Res. 10:1757-1771(2000).
RN   [6]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=C57BL/6J; TISSUE=Testis;
RA   Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA   Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA   Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA   Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA   Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA   Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA   Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA   Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA   Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA   Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA   Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK077224; BAC36695.1; -; mRNA.
DR   Ensembl; ENSMUSG00000049008; Mus musculus.
KW   Hypothetical protein.
SQ   SEQUENCE   389 AA;  40398 MW;  F8B69E073B81A82E CRC64;

  Query Match             38.9%;  Score 47.5;  DB 2;  Length 389;
  Best Local Similarity   38.5%;  Pred. No. 18;
  Matches   10;  Conservative    0;  Mismatches   15;  Indels    1;  Gaps    1;

Qy          3 CXXGPXTWXCXPYXCXXGPXTWXCXP 28
              |   | |  | |      | || | |
Db        230 CAEAP-TPVCTPPKLCAEPPTWVCTP 254


RESULT 14
Q4I355_GIBZE
ID   Q4I355_GIBZE PRELIMINARY;      PRT;   180 AA.
AC   Q4I355;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Predicted protein.
GN   ORFNames=FG08353.1;
OS   Gibberella zeae PH-1.
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC   Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
OX   NCBI_TaxID=229533;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=PH-1;
RA   Birren B., Nusbaum C., Abouelleil A., Allen N., Anderson S.,
RA   Arachchi H.M., Barna N., Bastien V., Bloom T., Boguslavkiy L.,
RA   Boukhgalter B., Butler J., Calvo S.E., Camarata J., Chang J.,
RA   Choepel Y., Collymore A., Cook A., Cooke P., Corum B., DeArellano K.,
RA   Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
RA   Erickson J., Faro S., Ferreira P., FitzGerald M., Gage D., Galagan J.,
RA   Gardyna S., Gnerre S., Graham L., Grand-Pierre N., Hafez N.,
RA   Hagopian D., Hagos B., Hall J., Horton L., Hulme W., Iliev I.,
RA   Jaffe D., Johnson R., Jones C., Kamal M., Kamat A., Karatas A.,
RA   Kells C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
RA   Ma L.-J., Mabbitt R., MacLean C., Macdonald P., Major J., Manning J.,
RA   Matthews C., Mauceli E., McCarthy M., Meldrim J., Meneus L.,
RA   Mihova T., Mlenga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
RA   Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neil D.,
RA   Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,
RA   Rachupka A., Ramasamy U., Raymond C., Retta R., Rise C., Rogov P.,
RA   Roman J., Schauer S., Schupback R., Seaman S., Severy P., Smirnov S.,
RA   Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,
RA   Talamas J., Tesfaye S., Theodore J., Topham K., Travers M.,
RA   Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,
RA   Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
RA   Lander E.;
RT   "Fusarium graminearum genome sequence.";
RL   Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; AACM01000335; EAA72141.1; -; Genomic_DNA.
SQ   SEQUENCE   180 AA;  20463 MW;  94C7B5242FEE6ED9 CRC64;

  Query Match             38.5%;  Score 47;  DB 2;  Length 180;
  Best Local Similarity   46.2%;  Pred. No. 10;
  Matches    6;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy         14 PYXCXXGPXTWXC 26
              |: |  ||  | |
Db         98 PHNCSFGPAPWEC 110


RESULT 15
Q9YK68_9HEPC
ID   Q9YK68_9HEPC PRELIMINARY;      PRT;   137 AA.
AC   Q9YK68;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Polyprotein (Fragment).
OS   Hepatitis C virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=98411414; PubMed=9738054;
RA   Wang Y.M., Ray S.C., Laeyendecker O., Ticehurst J.R., Thomas D.L.;
RT   "Assessment of hepatitis C virus sequence complexity by
RT   electrophoretic mobilities of both single-and double-stranded DNAs.";
RL   J. Clin. Microbiol. 36:2982-2989(1998).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RA   Wang Y.-M., Ray S.C., Laeyendecker O.B., Ticehurst J.R., Thomas D.L.;
RL   Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF073062; AAC61362.1; -; Genomic_RNA.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   InterPro; IPR002531; HCV_NS1.
DR   Pfam; PF01560; HCV_NS1; 1.
KW   Envelope protein; Polyprotein; Transmembrane.
```

```
FT   CHAIN        <1     14       E1.
FT   CHAIN        15   >137       E2.
FT   NON_TER       1      1
FT   NON_TER     137    137
SQ   SEQUENCE   137 AA;  14593 MW;  CFE4DFD9D3A6BEC1 CRC64;

  Query Match             37.7%;  Score 46;  DB 2;  Length 137;
  Best Local Similarity   36.4%;  Pred. No. 11;
  Matches    8;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy          7 PXTWXCXPYXCXXGPXTWXCXP 28
              | | | | |  |    | | |
Db        115 PYCWHCPPRPCGIVPAKSVCGP 136


Search completed: March 31, 2006, 16:35:08
Job time : 107.896 secs
```

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: March 31, 2006, 16:09:06 ; Search time 107.96 Seconds
(without alignments)
113.955 Million cell updates/sec

Title: US-10-609-217-84
Perfect score: 122
Sequence: 1 YXCXXGPXTWXCXPYXCXXGPXTWXCXP 28

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	110	90.2	28	3	AAB17028	Aab17028 EPO-mimet
2	110	90.2	28	5	ABB72811	Abb72811 Erythropo
3	110	90.2	28	7	ADJ72550	Adj72550 EPO mimet
4	110	90.2	28	8	ADJ52186	Adj52186 CH1 delet
5	110	90.2	28	8	ADJ51148	Adj51148 CH1 delet
6	99.5	81.6	29	3	AAB17029	Aab17029 EPO-mimet
7	99.5	81.6	29	5	ABB72812	Abb72812 Erythropo
8	99.5	81.6	29	7	ADJ72551	Adj72551 EPO mimet
9	99.5	81.6	29	8	ADJ52187	Adj52187 CH1 delet
10	99.5	81.6	29	8	ADJ51149	Adj51149 CH1 delet
11	99	81.1	40	3	AAB17036	Aab17036 EPO-mimet
12	99	81.1	40	5	ABB72819	Abb72819 Erythropo
13	99	81.1	40	8	ADJ52195	Adj52195 CH1 delet
14	98.5	80.7	41	3	AAB17037	Aab17037 EPO-mimet
15	98.5	80.7	41	5	ABB72820	Abb72820 Erythropo
16	98.5	80.7	41	7	ADJ72559	Adj72559 EPO mimet
17	98.5	80.7	41	8	ADJ51157	Adj51157 CH1 delet
18	97.5	79.9	46	3	AAB17039	Aab17039 EPO-mimet
19	97.5	79.9	46	5	ABB72822	Abb72822 Erythropo
20	97.5	79.9	70	7	ADJ72562	Adj72562 EPO mimet
21	97	79.5	47	3	AAB17040	Aab17040 EPO-mimet
22	97	79.5	47	8	ADJ52198	Adj52198 CH1 delet
23	97	79.5	47	8	ADJ51160	Adj51160 CH1 delet
24	97	79.5	49	5	ABB73393	Abb73393 EPO-mimet
25	97	79.5	49	5	ABB73392	Abb73392 EPO-mimet
26	97	79.5	50	3	AAB17283	Aab17283 EPO-mimet
27	97	79.5	50	3	AAB17284	Aab17284 EPO-mimet
28	97	79.5	57	3	AAB17314	Aab17314 EMP-EMP-F
29	97	79.5	57	5	ABB73408	Abb73408 EMP-EMP g
30	97	79.5	277	3	AAB16967	Aab16967 Fc-EMP-EM
31	97	79.5	277	3	AAB16966	Aab16966 EMP-EMP-F
32	97	79.5	277	5	ABB73418	Abb73418 Fc-EMP-EM
33	97	79.5	278	5	ABB73417	Abb73417 EMP-EMP-F
34	83	68.0	47	5	ABB72823	Abb72823 Erythropo
35	67	54.9	145	7	ADJ73529	Adj73529 Erythropo
36	63	51.6	20	2	AAY26373	Aay26373 Erythropo
37	63	51.6	20	2	AAY13719	Aay13719 Erythropo
38	63	51.6	20	2	AAY26402	Aay26402 Erythropo
39	63	51.6	20	2	AAY13665	Aay13665 Erythropo
40	63	51.6	20	2	AAW26979	Aaw26979 Monomer s
41	63	51.6	20	2	AAW27033	Aaw27033 Monomer s
42	63	51.6	23	2	AAY26392	Aay26392 Erythropo
43	62	50.8	20	2	AAY13713	Aay13713 Erythropo
44	62	50.8	20	2	AAY26360	Aay26360 Erythropo
45	62	50.8	20	2	AAW27027	Aaw27027 Monomer s

ALIGNMENTS

RESULT 1
AAB17028
ID AAB17028 standard; peptide; 28 AA.
XX
AC AAB17028;
XX
DT 31-OCT-2000 (first entry)
XX
DE EPO-mimetic peptide sequence SEQ ID NO:84.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US025044.
XX
PR 23-OCT-1998; 98US-0105371P.
PR 22-OCT-1999; 99US-00428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and pharmacologically
PT active peptides, useful for treating cancer and autoimmune diseases.
XX
PS Claim 13; Page 223; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,

CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and
CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC AAB18003 represent nucleotide and amino acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 28 AA;

Query Match 90.2%; Score 110; DB 3; Length 28;
Best Local Similarity 100.0%; Pred. No. 3.7e-09;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YXCXXGPXTWXCXPYXCXXGPXTWXCXP 28
||||||||||||||||||||||||||||
Db 1 YXCXXGPXTWXCXPYXCXXGPXTWXCXP 28

RESULT 2
ABB72811
ID ABB72811 standard; peptide; 28 AA.
XX
AC ABB72811;
XX
DT 05-APR-2002 (first entry)
XX
DE Erythropoietin (EPO) mimetic peptide SEQ ID NO:84.
XX
KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200183525-A2.
XX
PD 08-NOV-2001.
XX
PF 02-MAY-2001; 2001WO-US014310.
XX
PR 03-MAY-2000; 2000US-00563286.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
XX
DR WPI; 2002-130313/17.
XX
PT Novel vehicle-peptide molecule or its multimers useful for treating
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT diabetic retinopathy, obesity, sleep disorders and infertility.
XX
PS Claim 39; Page 41; 176pp; English.
XX
CC The present invention describes a vehicle-peptide molecule (I) or its
CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
CC neuroprotective activities. (I) can be used as a therapeutic or
CC prophylactic agent as well as for screening purposes. (I) is useful for
CC diagnosing diseases characterised by dysfunction of their associated
CC protein of interest, for identifying normal or abnormal proteins of
CC interest, as a part of diagnostic kit to detect the presence of their
CC proteins of interest in a biological sample. Additionally, (I) is useful
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
CC infertility, and neurological degenerative diseases. (I), comprising EPO-
CC mimetic compounds are useful for treating disorders characterised by low
CC red blood cell levels such as anaemia. The TPO-mimetic comprising
CC compounds are useful for treating conditions that involve an existing
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC represent amino acid and nucleic acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 28 AA;

Query Match 90.2%; Score 110; DB 5; Length 28;
Best Local Similarity 100.0%; Pred. No. 3.7e-09;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YXCXXGPXTWXCXPYXCXXGPXTWXCXP 28
||||||||||||||||||||||||||||
Db 1 YXCXXGPXTWXCXPYXCXXGPXTWXCXP 28

RESULT 3
ADJ72550
ID ADJ72550 standard; peptide; 28 AA.
XX
AC ADJ72550;
XX
DT 06-MAY-2004 (first entry)
XX
DE EPO mimetic peptide sequence SeqID 2.
XX
KW mimetic; CDR mimetibody; gene therapy; transgenic; immune;
KW cardiovascular; infectious; malignant; neurologic disease; anaemia;
KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;
KW erythropoietin; EPO.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 1. .28
FT /label= Xaa
FT /note= "Xaa can be any amino acid"
XX
PN WO2003084477-A2.
XX
PD 16-OCT-2003.
XX
PF 24-MAR-2003; 2003WO-US009139.
XX
PR 29-MAR-2002; 2002US-0368791P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Scallon BJ, Ghrayeb J;
XX
DR WPI; 2003-804237/75.
XX
PT New CDR mimetibody comprising a portion of a heavy or light chain
PT variable region comprising human framework or ligand binding region,
PT useful for preparing a composition for treating e.g., immune,
PT cardiovascular or neurologic disease.
XX
PS Disclosure; SEQ ID NO 2; 97pp; English.
XX

CC This invention relates to novel mammalian CDR mimetibodies, specific
CC portions or variants thereof. Specifically, it refers to an antibody
CC fragment where a protein has been inserted into, or replaces a portion
CC of, one or more CDR regions, such that each CDR mimetibody comprises at
CC least one portion of a heavy chain or light chain variable region, which
CC itself comprises at least one human framework region and at least one
CC ligand binding region (LBR). The present invention describes human
CC mimetibodies, including modified immunoglobulins and cleavage products
CC that can be useful in gene therapy and the generation of transgenic
CC plants and animals. Furthermore, the CDR mimetibody is useful for
CC preparing compositions for modulating, treating or reducing the symptoms
CC of immune, cardiovascular, infectious, malignant and/ or neurologic
CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,
CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This
CC peptide sequence is a erythropoietin (EPO) mimetic peptide sequence used
CC to make a mimetibody of the invention.
XX
SQ Sequence 28 AA;

Query Match 90.2%; Score 110; DB 7; Length 28;
Best Local Similarity 100.0%; Pred. No. 3.7e-09;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YXCXXGPXTWXCXPYXCXXGPXTWXCXP 28
||||||||||||||||||||||||||||
Db 1 YXCXXGPXTWXCXPYXCXXGPXTWXCXP 28

RESULT 4
ADJ52186
ID ADJ52186 standard; peptide; 28 AA.
XX
AC ADJ52186;
XX
DT 06-MAY-2004 (first entry)
XX
DE CH1 deleted mimetibody-related peptide SeqID2.
XX
KW CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;
KW fungicide; gene therapy; immune disorder; cardiovascular disease;
KW arrhythmia; hypertension; heart failure; neurodegenerative;
KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;
KW cancerous condition; infectious disease; bacterial infection;
KW viral infection; fungal infection.
XX
OS Unidentified.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 1. .28
FT /note= "All Xaa's in this sequence are unidentified amino
FT acids"
XX
PN WO2004002417-A2.
XX
PD 08-JAN-2004.
XX
PF 27-JUN-2003; 2003WO-US020347.
XX
PR 28-JUN-2002; 2002US-0392431P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Ghrayeb J, Scallon BJ, Nesspor TC;
PI Kutoloski KA;
XX
DR WPI; 2004-082870/08.
XX
PT New CH1-deleted mimetibody polypeptides and nucleic acids, useful for
PT modulating, treating, alleviating, preventing an immune, cardiovascular,
PT or neurodegenerative disease or disorder, anemia, cancer, or infectious
PT diseases.
XX
PS Disclosure; SEQ ID NO 2; 129pp; English.
XX
CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an immunosuppressive,
CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,
CC antibacterial, virucide or fungicide activity. In addition, the disclosed
CC sequences may prove useful for gene therapy. The CH1-deleted mimetibody
CC is useful for diagnosing or treating a disease condition in a cell,
CC tissue, organ or animal, specifically for modulating, treating,
CC alleviating, preventing the incidence or reducing the symptoms of an
CC immune, cardiovascular (for example arrhythmia, hypertension or heart
CC failure), or neurodegenerative (for example multiple sclerosis, dementia
CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous
CC conditions, or infectious diseases (for example bacterial, viral or
CC fungal infection). The present sequence is that of a peptide which may be
CC used during the creation of a mimetibody of the invention.
XX
SQ Sequence 28 AA;

Query Match 90.2%; Score 110; DB 8; Length 28;
Best Local Similarity 100.0%; Pred. No. 3.7e-09;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YXCXXGPXTWXCXPYXCXXGPXTWXCXP 28
||||||||||||||||||||||||||||
Db 1 YXCXXGPXTWXCXPYXCXXGPXTWXCXP 28

RESULT 5
ADJ51148
ID ADJ51148 standard; peptide; 28 AA.
XX
AC ADJ51148;
XX
DT 06-MAY-2004 (first entry)
XX
DE CH1 deleted mimetibody-related peptide SeqID2.
XX
KW CH1 deleted mimetibody; osteopathic; cardiovascular-Gen;
KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;
KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;
KW antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;
KW ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;
KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;
KW dental disorder; oral disorder; dermatological disorder; ear disorder;
KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;
KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;
KW obstetric disorder; haematologic disorder; immunological disorder;
KW allergic disorder; infectious disorder; musculoskeletal disorder;
KW oncological disorder; neurological disorder; nutritional disorder;
KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;
KW renal disorder; pulmonary disorder.
XX
OS Unidentified.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 1. .28
FT /note= "All Xaa's in this sequence are unidentified amino
FT acids"
XX
PN WO2004002424-A2.
XX
PD 08-JAN-2004.
XX
PF 30-JUN-2003; 2003WO-US020495.
XX
PR 28-JUN-2002; 2002US-0392431P.
PR 19-SEP-2002; 2002US-0412144P.
XX
PA (CENZ) CENTOCOR INC.

XX
PI Heavner GA, Knight DM, Ghrayeb J, Scallon BJ, Nesspor TC;
PI Kutoloski KA;
XX
DR WPI; 2004-082872/08.
XX
PT New CH1 deleted mimetibody polypeptide and nucleic acid, useful for
PT diagnosing, preventing or treating cardiovascular, dermatologic,
PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and
PT nutritional disorders.
XX
PS Claim 8; SEQ ID NO 2; 123pp; English.
XX
CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an osteopathic,
CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,
CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,
CC immunomodulator, antiallergic, muscular-Gen, cytostatic,
CC antiinflammatory, neuroleptic, ophthalmological, nephrotropic or
CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-
CC modulator or cytokine-agonist. The methods and compositions of the
CC present invention are useful for the diagnosis, prevention and/or
CC treatment of diseases or conditions associated with aberrant expression
CC or activity of the CH1 deleted mimetibody, such as a bone or joint,
CC cardiovascular, dental or oral, dermatological, ear, nose or throat,
CC endocrine, metabolic, gastrointestinal, gynaecological, hepatic,
CC obstetric, haematologic, immunological, allergic, infectious,
CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,
CC pediatric, psychiatric, renal or pulmonary disorders. The present
CC sequence is that of a peptide which may be used during the creation of a
CC mimetibody of the invention.
XX
SQ Sequence 28 AA;

Query Match 90.2%; Score 110; DB 8; Length 28;
Best Local Similarity 100.0%; Pred. No. 3.7e-09;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YXCXXGPXTWXCXPYXCXXGPXTWXCXP 28
||||||||||||||||||||||||||||
Db 1 YXCXXGPXTWXCXPYXCXXGPXTWXCXP 28

RESULT 6
AAB17029
ID AAB17029 standard; peptide; 29 AA.
XX
AC AAB17029;
XX
DT 31-OCT-2000 (first entry)
XX
DE EPO-mimetic peptide sequence SEQ ID NO:85.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US025044.
XX
PR 23-OCT-1998; 98US-0105371P.
PR 22-OCT-1999; 99US-00428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and pharmacologically
PT active peptides, useful for treating cancer and autoimmune diseases.
XX
PS Claim 13; Page 224; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and
CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC AAB18003 represent nucleotide and amino acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 29 AA;

Query Match 81.6%; Score 99.5; DB 3; Length 29;
Best Local Similarity 96.6%; Pred. No. 1.3e-07;
Matches 28; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 1 YXCXXGPXTWXCXP-YXCXXGPXTWXCXP 28
||||||||||||||| ||||||||||||||
Db 1 YXCXXGPXTWXCXPXYXCXXGPXTWXCXP 29

RESULT 7
ABB72812
ID ABB72812 standard; peptide; 29 AA.
XX
AC ABB72812;
XX
DT 05-APR-2002 (first entry)
XX
DE Erythropoietin (EPO) mimetic peptide SEQ ID NO:85.
XX
KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200183525-A2.
XX
PD 08-NOV-2001.
XX
PF 02-MAY-2001; 2001WO-US014310.
XX

PR 03-MAY-2000; 2000US-00563286.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
XX
DR WPI; 2002-130313/17.
XX
PT Novel vehicle-peptide molecule or its multimers useful for treating
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT diabetic retinopathy, obesity, sleep disorders and infertility.
XX
PS Claim 39; Page 41; 176pp; English.
XX
CC The present invention describes a vehicle-peptide molecule (I) or its
CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
CC neuroprotective activities. (I) can be used as a therapeutic or
CC prophylactic agent as well as for screening purposes. (I) is useful for
CC diagnosing diseases characterised by dysfunction of their associated
CC protein of interest, for identifying normal or abnormal proteins of
CC interest, as a part of diagnostic kit to detect the presence of their
CC proteins of interest in a biological sample. Additionally, (I) is useful
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
CC infertility, and neurological degenerative diseases. (I), comprising EPO-
CC mimetic compounds are useful for treating disorders characterised by low
CC red blood cell levels such as anaemia. The TPO-mimetic comprising
CC compounds are useful for treating conditions that involve an existing
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC represent amino acid and nucleic acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 29 AA;

Query Match 81.6%; Score 99.5; DB 5; Length 29;
Best Local Similarity 96.6%; Pred. No. 1.3e-07;
Matches 28; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 1 YXCXXGPXTWXCXP-YXCXXGPXTWXCXP 28
||||||||||||||| ||||||||||||||
Db 1 YXCXXGPXTWXCXPXYXCXXGPXTWXCXP 29

RESULT 8
ADJ72551
ID ADJ72551 standard; peptide; 29 AA.
XX
AC ADJ72551;
XX
DT 06-MAY-2004 (first entry)
XX
DE EPO mimetic peptide sequence SeqID 3.
XX
KW mimetic; CDR mimetibody; gene therapy; transgenic; immune;
KW cardiovascular; infectious; malignant; neurologic disease; anaemia;
KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;
KW erythropoietin; EPO.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 1. .29
FT /label= Xaa
FT /note= "Xaa can be any amino acid"
XX
PN WO2003084477-A2.
XX
PD 16-OCT-2003.
XX
PF 24-MAR-2003; 2003WO-US009139.
XX
PR 29-MAR-2002; 2002US-0368791P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Scallon BJ, Ghrayeb J;
XX
DR WPI; 2003-804237/75.
XX
PT New CDR mimetibody comprising a portion of a heavy or light chain
PT variable region comprising human framework or ligand binding region,
PT useful for preparing a composition for treating e.g., immune,
PT cardiovascular or neurologic disease.
XX
PS Disclosure; SEQ ID NO 3; 97pp; English.
XX
CC This invention relates to novel mammalian CDR mimetibodies, specific
CC portions or variants thereof. Specifically, it refers to an antibody
CC fragment where a protein has been inserted into, or replaces a portion
CC of, one or more CDR regions, such that each CDR mimetibody comprises at
CC least one portion of a heavy chain or light chain variable region, which
CC itself comprises at least one human framework region and at least one
CC ligand binding region (LBR). The present invention describes human
CC mimetibodies, including modified immunoglobulins and cleavage products
CC that can be useful in gene therapy and the generation of transgenic
CC plants and animals. Furthermore, the CDR mimetibody is useful for
CC preparing compositions for modulating, treating or reducing the symptoms
CC of immune, cardiovascular, infectious, malignant and/ or neurologic
CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,
CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This
CC peptide sequence is a erythropoietin (EPO) mimetic peptide sequence used
CC to make a mimetibody of the invention.
XX
SQ Sequence 29 AA;

Query Match 81.6%; Score 99.5; DB 7; Length 29;
Best Local Similarity 96.6%; Pred. No. 1.3e-07;
Matches 28; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 1 YXCXXGPXTWXCXP-YXCXXGPXTWXCXP 28
||||||||||||||| ||||||||||||||
Db 1 YXCXXGPXTWXCXPAYXCXXGPXTWXCXP 29

RESULT 9
ADJ52187
ID ADJ52187 standard; peptide; 29 AA.
XX
AC ADJ52187;
XX
DT 06-MAY-2004 (first entry)
XX
DE CH1 deleted mimetibody-related peptide SeqID3.
XX
KW CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;
KW fungicide; gene therapy; immune disorder; cardiovascular disease;
KW arrhythmia; hypertension; heart failure; neurodegenerative;
KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;
KW cancerous condition; infectious disease; bacterial infection;
KW viral infection; fungal infection.
XX
OS Unidentified.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 1. .29
FT /note= "All Xaa's in this sequence are unidentified amino
FT acids"
XX
PN WO2004002417-A2.

XX
PD 08-JAN-2004.
XX
PF 27-JUN-2003; 2003WO-US020347.
XX
PR 28-JUN-2002; 2002US-0392431P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Ghrayeb J, Scallon BJ, Nesspor TC;
PI Kutoloski KA;
XX
DR WPI; 2004-082870/08.
XX
PT New CH1-deleted mimetibody polypeptides and nucleic acids, useful for
PT modulating, treating, alleviating, preventing an immune, cardiovascular,
PT or neurodegenerative disease or disorder, anemia, cancer, or infectious
PT diseases.
XX
PS Disclosure; SEQ ID NO 3; 129pp; English.
XX
CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an immunosuppressive,
CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,
CC antibacterial, virucide or fungicide activity. In addition, the disclosed
CC sequences may prove useful for gene therapy. The CH1-deleted mimetibody
CC is useful for diagnosing or treating a disease condition in a cell,
CC tissue, organ or animal, specifically for modulating, treating,
CC alleviating, preventing the incidence or reducing the symptoms of an
CC immune, cardiovascular (for example arrhythmia, hypertension or heart
CC failure), or neurodegenerative (for example multiple sclerosis, dementia
CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous
CC conditions, or infectious diseases (for example bacterial, viral or
CC fungal infection). The present sequence is that of a peptide which may be
CC used during the creation of a mimetibody of the invention.
XX
SQ Sequence 29 AA;

Query Match 81.6%; Score 99.5; DB 8; Length 29;
Best Local Similarity 96.6%; Pred. No. 1.3e-07;
Matches 28; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 1 YXCXXGPXTWXCXP-YXCXXGPXTWXCXP 28
||||||||||||||| |||||||||||||
Db 1 YXCXXGPXTWXCXPAYXCXXGPXTWXCXP 29

RESULT 10
ADJ51149
ID ADJ51149 standard; peptide; 29 AA.
XX
AC ADJ51149;
XX
DT 06-MAY-2004 (first entry)
XX
DE CH1 deleted mimetibody-related peptide SeqID3.
XX
KW CH1 deleted mimetibody; osteopathic; cardiovascular-Gen;
KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;
KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;
KW antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;
KW ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;
KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;
KW dental disorder; oral disorder; dermatological disorder; ear disorder;
KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;
KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;
KW obstetric disorder; haematologic disorder; immunological disorder;
KW allergic disorder; infectious disorder; musculoskeletal disorder;
KW oncological disorder; neurological disorder; nutritional disorder;
KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;
KW renal disorder; pulmonary disorder.
XX
OS Unidentified.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 1. .29
FT /note= "All Xaa's in this sequence are unidentified amino
FT acids"
XX
PN WO2004002424-A2.
XX
PD 08-JAN-2004.
XX
PF 30-JUN-2003; 2003WO-US020495.
XX
PR 28-JUN-2002; 2002US-0392431P.
PR 19-SEP-2002; 2002US-0412144P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Ghrayeb J, Scallon BJ, Nesspor TC;
PI Kutoloski KA;
XX
DR WPI; 2004-082872/08.
XX
PT New CH1 deleted mimetibody polypeptide and nucleic acid, useful for
PT diagnosing, preventing or treating cardiovascular, dermatologic,
PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and
PT nutritional disorders.
XX
PS Claim 8; SEQ ID NO 3; 123pp; English.
XX
CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an osteopathic,
CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,
CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,
CC immunomodulator, antiallergic, muscular-Gen, cytostatic,
CC antiinflammatory, neuroleptic, ophthalmological, nephrotropic or
CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-
CC modulator or cytokine-agonist. The methods and compositions of the
CC present invention are useful for the diagnosis, prevention and/or
CC treatment of diseases or conditions associated with aberrant expression
CC or activity of the CH1 deleted mimetibody, such as a bone or joint,
CC cardiovascular, dental or oral, dermatological, ear, nose or throat,
CC endocrine, metabolic, gastrointestinal, gynaecological, hepatic,
CC obstetric, haematologic, immunological, allergic, infectious,
CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,
CC pediatric, psychiatric, renal or pulmonary disorders. The present
CC sequence is that of a peptide which may be used during the creation of a
CC mimetibody of the invention.
XX
SQ Sequence 29 AA;

Query Match 81.6%; Score 99.5; DB 8; Length 29;
Best Local Similarity 96.6%; Pred. No. 1.3e-07;
Matches 28; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 1 YXCXXGPXTWXCXP-YXCXXGPXTWXCXP 28
||||||||||||||| |||||||||||||
Db 1 YXCXXGPXTWXCXPAYXCXXGPXTWXCXP 29

RESULT 11
AAB17036
ID AAB17036 standard; peptide; 40 AA.
XX
AC AAB17036;
XX
DT 31-OCT-2000 (first entry)
XX
DE EPO-mimetic peptide sequence SEQ ID NO:92.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;

KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US025044.
XX
PR 23-OCT-1998; 98US-0105371P.
PR 22-OCT-1999; 99US-00428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and pharmacologically
PT active peptides, useful for treating cancer and autoimmune diseases.
XX
PS Claim 13; Page 226; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and
CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC AAB18003 represent nucleotide and amino acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 40 AA;

Query Match 81.1%; Score 99; DB 3; Length 40;
Best Local Similarity 47.1%; Pred. No. 2.1e-07;
Matches 16; Conservative 0; Mismatches 12; Indels 6; Gaps 1;

Qy 1 YXCXXGPXTWXCXP------YXCXXGPXTWXCXP 28
Db 4 YSCHFGPLTWVCKPQGGGGTYSCHFGPLTWVCKP 37

RESULT 12
ABB72819
ID ABB72819 standard; peptide; 40 AA.
XX
AC ABB72819;
XX
DT 05-APR-2002 (first entry)
XX
DE Erythropoietin (EPO) mimetic peptide SEQ ID NO:92.
XX
KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200183525-A2.
XX
PD 08-NOV-2001.
XX
PF 02-MAY-2001; 2001WO-US014310.
XX
PR 03-MAY-2000; 2000US-00563286.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
XX
DR WPI; 2002-130313/17.
XX
PT Novel vehicle-peptide molecule or its multimers useful for treating
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT diabetic retinopathy, obesity, sleep disorders and infertility.
XX
PS Claim 39; Page 41; 176pp; English.
XX
CC The present invention describes a vehicle-peptide molecule (I) or its
CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
CC neuroprotective activities. (I) can be used as a therapeutic or
CC prophylactic agent as well as for screening purposes. (I) is useful for
CC diagnosing diseases characterised by dysfunction of their associated
CC protein of interest, for identifying normal or abnormal proteins of
CC interest, as a part of diagnostic kit to detect the presence of their
CC proteins of interest in a biological sample. Additionally, (I) is useful
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
CC infertility, and neurological degenerative diseases. (I), comprising EPO-
CC mimetic compounds are useful for treating disorders characterised by low
CC red blood cell levels such as anaemia. The TPO-mimetic comprising
CC compounds are useful for treating conditions that involve an existing
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC represent amino acid and nucleic acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 40 AA;

Query Match 81.1%; Score 99; DB 5; Length 40;
Best Local Similarity 47.1%; Pred. No. 2.1e-07;
Matches 16; Conservative 0; Mismatches 12; Indels 6; Gaps 1;

Qy 1 YXCXXGPXTWXCXP------YXCXXGPXTWXCXP 28
Db 4 YSCHFGPLTWVCKPQGGGGTYSCHFGPLTWVCKP 37

RESULT 13
ADJ52195
ID ADJ52195 standard; peptide; 40 AA.
XX
AC ADJ52195;
XX
DT 06-MAY-2004 (first entry)

XX
DE CH1 deleted mimetibody-related peptide SeqID11.
XX
KW CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;
KW fungicide; gene therapy; immune disorder; cardiovascular disease;
KW arrhythmia; hypertension; heart failure; neurodegenerative;
KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;
KW cancerous condition; infectious disease; bacterial infection;
KW viral infection; fungal infection.
XX
OS Unidentified.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 21
FT /label= OTHER
FT /note= "OTHER= linker"
XX
PN WO2004002417-A2.
XX
PD 08-JAN-2004.
XX
PF 27-JUN-2003; 2003WO-US020347.
XX
PR 28-JUN-2002; 2002US-0392431P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Ghrayeb J, Scallon BJ, Nesspor TC;
PI Kutoloski KA;
XX
DR WPI; 2004-082870/08.
XX
PT New CH1-deleted mimetibody polypeptides and nucleic acids, useful for
PT modulating, treating, alleviating, preventing an immune, cardiovascular,
PT or neurodegenerative disease or disorder, anemia, cancer, or infectious
PT diseases.
XX
PS Disclosure; SEQ ID NO 11; 129pp; English.
XX
CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an immunosuppressive,
CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,
CC antibacterial, virucide or fungicide activity. In addition, the disclosed
CC sequences may prove useful for gene therapy. The CH1-deleted mimetibody
CC is useful for diagnosing or treating a disease condition in a cell,
CC tissue, organ or animal, specifically for modulating, treating,
CC alleviating, preventing the incidence or reducing the symptoms of an
CC immune, cardiovascular (for example arrhythmia, hypertension or heart
CC failure), or neurodegenerative (for example multiple sclerosis, dementia
CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous
CC conditions, or infectious diseases (for example bacterial, viral or
CC fungal infection). The present sequence is that of a peptide which may be
CC used during the creation of a mimetibody of the invention.
XX
SQ Sequence 40 AA;

Query Match 81.1%; Score 99; DB 8; Length 40;
Best Local Similarity 47.1%; Pred. No. 2.1e-07;
Matches 16; Conservative 0; Mismatches 12; Indels 6; Gaps 1;

Qy 1 YXCXXGPXTWXCXP------YXCXXGPXTWXCXP 28
Db 4 YSCHFGPLTWVCKPQGGGGTYSCHFGPLTWVCKP 37

RESULT 14
AAB17037
ID AAB17037 standard; peptide; 41 AA.
XX
AC AAB17037;
XX
DT 31-OCT-2000 (first entry)
XX
DE EPO-mimetic peptide sequence SEQ ID NO:93.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US025044.
XX
PR 23-OCT-1998; 98US-0105371P.
PR 22-OCT-1999; 99US-00428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and pharmacologically
PT active peptides, useful for treating cancer and autoimmune diseases.
XX
PS Claim 13; Page 227; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and
CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC AAB18003 represent nucleotide and amino acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 41 AA;

Query Match 80.7%; Score 98.5; DB 3; Length 41;
Best Local Similarity 45.7%; Pred. No. 2.5e-07;
Matches 16; Conservative 0; Mismatches 12; Indels 7; Gaps 1;

Qy 1 YXCXXGPXTWXCXP-------YXCXXGPXTWXCXP 28
Db 4 YSCHFGPLTWVCKPQGGXGGTYSCHFGPLTWVCKP 38

RESULT 15
ABB72820
ID ABB72820 standard; peptide; 41 AA.
XX
AC ABB72820;
XX
DT 05-APR-2002 (first entry)
XX

DE Erythropoietin (EPO) mimetic peptide SEQ ID NO:93.
XX
KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNP-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200183525-A2.
XX
PD 08-NOV-2001.
XX
PF 02-MAY-2001; 2001WO-US014310.
XX
PR 03-MAY-2000; 2000US-00563286.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
XX
DR WPI; 2002-130313/17.
XX
PT Novel vehicle-peptide molecule or its multimers useful for treating
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT diabetic retinopathy, obesity, sleep disorders and infertility.
XX
PS Claim 39; Page 41; 176pp; English.
XX
CC The present invention describes a vehicle-peptide molecule (I) or its
CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
CC neuroprotective activities. (I) can be used as a therapeutic or
CC prophylactic agent as well as for screening purposes. (I) is useful for
CC diagnosing diseases characterised by dysfunction of their associated
CC protein of interest, for identifying normal or abnormal proteins of
CC interest, as a part of diagnostic kit to detect the presence of their
CC proteins of interest in a biological sample. Additionally, (I) is useful
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
CC infertility, and neurological degenerative diseases. (I), comprising EPO-
CC mimetic compounds are useful for treating disorders characterised by low
CC red blood cell levels such as anaemia. The TPO-mimetic comprising
CC compounds are useful for treating conditions that involve an existing
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC represent amino acid and nucleic acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 41 AA;

Query Match 80.7%; Score 98.5; DB 5; Length 41;
Best Local Similarity 45.7%; Pred. No. 2.5e-07;
Matches 16; Conservative 0; Mismatches 12; Indels 7; Gaps 1;

```
Qy          1 YXCXXGPXTWXCXP-------YXCXXGPXTWXCXP 28
              | |  || || | |       | |  || || | |
Db          4 YSCHFGPLTWVCKPQGGXGGTYSCHFGPLTWVCKP 38
```

Search completed: March 31, 2006, 16:22:26

Job time : 108.96 secs

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: March 31, 2006, 16:35:37 ; Search time 27.4428 Seconds
(without alignments)
84.354 Million cell updates/sec

Title: US-10-609-217-84
Perfect score: 122
Sequence: 1 YXCXXGPXTWXCXPYXCXXGPXTWXCXP 28

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/6_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/H_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/RE_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	110	90.2	28	2	US-09-428-082B-84	Sequence 84, Appl
2	99	81.1	40	2	US-09-428-082B-92	Sequence 92, Appl
3	97.5	79.9	46	2	US-09-428-082B-95	Sequence 95, Appl
4	97	79.5	49	2	US-09-428-082B-339	Sequence 339, App
5	97	79.5	49	2	US-09-428-082B-340	Sequence 340, App
6	97	79.5	57	2	US-09-428-082B-417	Sequence 417, App
7	97	79.5	277	2	US-09-428-082B-20	Sequence 20, Appl
8	97	79.5	277	2	US-09-428-082B-22	Sequence 22, Appl
9	63	51.6	20	1	US-08-484-135-24	Sequence 24, Appl
10	63	51.6	20	1	US-08-484-135-78	Sequence 78, Appl
11	63	51.6	20	1	US-08-484-635-40	Sequence 40, Appl
12	63	51.6	20	1	US-08-484-635-206	Sequence 206, App
13	63	51.6	20	1	US-08-484-631-40	Sequence 40, Appl
14	63	51.6	20	1	US-08-484-631-206	Sequence 206, App
15	63	51.6	20	1	US-08-827-570-40	Sequence 40, Appl
16	63	51.6	20	1	US-08-827-570-206	Sequence 206, App
17	63	51.6	23	1	US-08-484-635-56	Sequence 56, Appl
18	63	51.6	23	1	US-08-484-631-56	Sequence 56, Appl
19	63	51.6	23	1	US-08-827-570-56	Sequence 56, Appl
20	62	50.8	20	1	US-08-484-135-72	Sequence 72, Appl
21	62	50.8	20	1	US-08-484-635-30	Sequence 30, Appl
22	62	50.8	20	1	US-08-484-631-30	Sequence 30, Appl
23	62	50.8	20	1	US-08-827-570-30	Sequence 30, Appl
24	61	50.0	21	1	US-08-484-635-101	Sequence 101, App
25	61	50.0	21	1	US-08-484-631-101	Sequence 101, App
26	61	50.0	21	1	US-08-827-570-101	Sequence 101, App
27	59	48.4	20	1	US-08-484-635-60	Sequence 60, Appl
28	59	48.4	20	1	US-08-484-631-60	Sequence 60, Appl
29	59	48.4	20	1	US-08-827-570-60	Sequence 60, Appl
30	59	48.4	253	2	US-09-428-082B-18	Sequence 18, Appl
31	58	47.5	20	1	US-08-484-135-46	Sequence 46, Appl
32	58	47.5	20	1	US-08-484-635-219	Sequence 219, App
33	58	47.5	20	1	US-08-484-631-219	Sequence 219, App
34	58	47.5	20	1	US-08-827-570-219	Sequence 219, App
35	58	47.5	22	1	US-08-484-135-68	Sequence 68, Appl
36	58	47.5	22	1	US-08-484-635-25	Sequence 25, Appl
37	58	47.5	22	1	US-08-484-631-25	Sequence 25, Appl
38	58	47.5	22	1	US-08-827-570-25	Sequence 25, Appl
39	58	47.5	24	1	US-08-484-635-83	Sequence 83, Appl
40	58	47.5	24	1	US-08-484-631-83	Sequence 83, Appl
41	58	47.5	24	1	US-08-827-570-83	Sequence 83, Appl
42	57.5	47.1	39	2	US-09-428-082B-395	Sequence 395, App
43	57	46.7	20	1	US-08-484-135-11	Sequence 11, Appl
44	57	46.7	20	1	US-08-484-135-35	Sequence 35, Appl
45	57	46.7	20	1	US-08-484-135-87	Sequence 87, Appl

ALIGNMENTS

RESULT 1
US-09-428-082B-84
; Sequence 84, Application US/09428082B
; Patent No. 6660843
; GENERAL INFORMATION:
; APPLICANT: FEIGE, ULRICH
; APPLICANT: LIU, CHUAN-FA
; APPLICANT: CHEETHAM, JANET C.
; APPLICANT: BOONE, THOMAS CHARLES
; TITLE OF INVENTION: MODIFIED PEPTIDES AS THERAPEUTIC AGENTS
; FILE REFERENCE: A-527
; CURRENT APPLICATION NUMBER: US/09/428,082B
; CURRENT FILING DATE: 1999-10-22
; PRIOR APPLICATION NUMBER: 60/105,371
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 1133
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 84
; LENGTH: 28
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: EPO-MIMETIC PEPTIDE
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (2, 4, 5, 8, 11, 13, 16, 18, 19, 22, 25)..(27)
; OTHER INFORMATION: Xaa = any amino acid
US-09-428-082B-84

Query Match 90.2%; Score 110; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 3.1e-09;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YXCXXGPXTWXCXPYXCXXGPXTWXCXP 28
||||||||||||||||||||||||||||
Db 1 YXCXXGPXTWXCXPYXCXXGPXTWXCXP 28

RESULT 2
US-09-428-082B-92
; Sequence 92, Application US/09428082B
; Patent No. 6660843
; GENERAL INFORMATION:
; APPLICANT: FEIGE, ULRICH
; APPLICANT: LIU, CHUAN-FA
; APPLICANT: CHEETHAM, JANET C.
; APPLICANT: BOONE, THOMAS CHARLES
; TITLE OF INVENTION: MODIFIED PEPTIDES AS THERAPEUTIC AGENTS
; FILE REFERENCE: A-527
; CURRENT APPLICATION NUMBER: US/09/428,082B

```
;  CURRENT FILING DATE:  1999-10-22
;  PRIOR APPLICATION NUMBER: 60/105,371
;  PRIOR FILING DATE: 1998-10-23
;  NUMBER OF SEQ ID NOS: 1133
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 92
;   LENGTH: 40
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: EPO-MIMETIC PEPTIDE
US-09-428-082B-92

  Query Match             81.1%;  Score 99;  DB 2;  Length 40;
  Best Local Similarity   47.1%;  Pred. No. 1.5e-07;
  Matches   16;  Conservative    0;  Mismatches   12;  Indels    6;  Gaps    1;

Qy          1 YXCXXGPXTWXCXP------YXCXXGPXTWXCXP 28
              | |  || || | |      | |  || || | |
Db          4 YSCHFGPLTWVCKPQGGGGTYSCHFGPLTWVCKP 37


RESULT 3
US-09-428-082B-95
; Sequence 95, Application US/09428082B
; Patent No. 6660843
; GENERAL INFORMATION:
;  APPLICANT: FEIGE, ULRICH
;  APPLICANT:  LIU, CHUAN-FA
;  APPLICANT:  CHEETHAM, JANET C.
;  APPLICANT:  BOONE, THOMAS CHARLES
;  TITLE OF INVENTION: MODIFIED PEPTIDES AS THERAPEUTIC AGENTS
;  FILE REFERENCE: A-527
;  CURRENT APPLICATION NUMBER: US/09/428,082B
;  CURRENT FILING DATE:  1999-10-22
;  PRIOR APPLICATION NUMBER: 60/105,371
;  PRIOR FILING DATE: 1998-10-23
;  NUMBER OF SEQ ID NOS: 1133
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 95
;   LENGTH: 46
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: EPO-MIMETIC PEPTIDE
US-09-428-082B-95

  Query Match             79.9%;  Score 97.5;  DB 2;  Length 46;
  Best Local Similarity   43.2%;  Pred. No. 2.7e-07;
  Matches   16;  Conservative    0;  Mismatches   12;  Indels    9;  Gaps    1;

Qy          1 YXCXXGPXTWXCXP---------YXCXXGPXTWXCXP 28
              | |  || || | |         | |  || || | |
Db          4 YSCHFGPLTWVCKPQGGSSKGGTYSCHFGPLTWVCKP 40


RESULT 4
US-09-428-082B-339
; Sequence 339, Application US/09428082B
; Patent No. 6660843
; GENERAL INFORMATION:
;  APPLICANT: FEIGE, ULRICH
;  APPLICANT:  LIU, CHUAN-FA
;  APPLICANT:  CHEETHAM, JANET C.
;  APPLICANT:  BOONE, THOMAS CHARLES
;  TITLE OF INVENTION: MODIFIED PEPTIDES AS THERAPEUTIC AGENTS
;  FILE REFERENCE: A-527
;  CURRENT APPLICATION NUMBER: US/09/428,082B
;  CURRENT FILING DATE:  1999-10-22
;  PRIOR APPLICATION NUMBER: 60/105,371
;  PRIOR FILING DATE: 1998-10-23
;  NUMBER OF SEQ ID NOS: 1133
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 339
;   LENGTH: 49
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: EPO-MIMETIC
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(1)
;   OTHER INFORMATION: Fc domain attached at Position 1 of the N-terminus
US-09-428-082B-339

  Query Match             79.5%;  Score 97;  DB 2;  Length 49;
  Best Local Similarity   42.1%;  Pred. No. 3.4e-07;
  Matches   16;  Conservative    0;  Mismatches   12;  Indels   10;  Gaps    1;

Qy          1 YXCXXGPXTWXCXP----------YXCXXGPXTWXCXP 28
              | |  || || | |          | |  || || | |
Db          9 YSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKP 46


RESULT 5
US-09-428-082B-340
; Sequence 340, Application US/09428082B
; Patent No. 6660843
; GENERAL INFORMATION:
;  APPLICANT: FEIGE, ULRICH
;  APPLICANT:  LIU, CHUAN-FA
;  APPLICANT:  CHEETHAM, JANET C.
;  APPLICANT:  BOONE, THOMAS CHARLES
;  TITLE OF INVENTION: MODIFIED PEPTIDES AS THERAPEUTIC AGENTS
;  FILE REFERENCE: A-527
;  CURRENT APPLICATION NUMBER: US/09/428,082B
;  CURRENT FILING DATE:  1999-10-22
;  PRIOR APPLICATION NUMBER: 60/105,371
;  PRIOR FILING DATE: 1998-10-23
;  NUMBER OF SEQ ID NOS: 1133
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 340
;   LENGTH: 49
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: EPO-MIMETIC
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (49)..(49)
;   OTHER INFORMATION: Fc domain attached at Position 49 of the C-terminus
US-09-428-082B-340

  Query Match             79.5%;  Score 97;  DB 2;  Length 49;
  Best Local Similarity   42.1%;  Pred. No. 3.4e-07;
  Matches   16;  Conservative    0;  Mismatches   12;  Indels   10;  Gaps    1;

Qy          1 YXCXXGPXTWXCXP----------YXCXXGPXTWXCXP 28
              | |  || || | |          | |  || || | |
Db          4 YSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKP 41


RESULT 6
US-09-428-082B-417
; Sequence 417, Application US/09428082B
; Patent No. 6660843
; GENERAL INFORMATION:
;  APPLICANT: FEIGE, ULRICH
;  APPLICANT:  LIU, CHUAN-FA
;  APPLICANT:  CHEETHAM, JANET C.
;  APPLICANT:  BOONE, THOMAS CHARLES
;  TITLE OF INVENTION: MODIFIED PEPTIDES AS THERAPEUTIC AGENTS
;  FILE REFERENCE: A-527
;  CURRENT APPLICATION NUMBER: US/09/428,082B
;  CURRENT FILING DATE:  1999-10-22
;  PRIOR APPLICATION NUMBER: 60/105,371
```

```
;  PRIOR FILING DATE: 1998-10-23
;  NUMBER OF SEQ ID NOS: 1133
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 417
;   LENGTH: 57
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: EMP-EMP-Fc
US-09-428-082B-417

  Query Match             79.5%;  Score 97;  DB 2;  Length 57;
  Best Local Similarity   42.1%;  Pred. No. 3.9e-07;
  Matches   16;  Conservative    0;  Mismatches   12;  Indels   10;  Gaps    1;

Qy          1 YXCXXGPXTWXCXP----------YXCXXGPXTWXCXP 28
              | |  || || | |           | |  || || | |
Db          5 YSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKP 42


RESULT 7
US-09-428-082B-20
; Sequence 20, Application US/09428082B
; Patent No. 6660843
; GENERAL INFORMATION:
;  APPLICANT: FEIGE, ULRICH
;  APPLICANT:  LIU, CHUAN-FA
;  APPLICANT:  CHEETHAM, JANET C.
;  APPLICANT:  BOONE, THOMAS CHARLES
;  TITLE OF INVENTION: MODIFIED PEPTIDES AS THERAPEUTIC AGENTS
;  FILE REFERENCE: A-527
;  CURRENT APPLICATION NUMBER: US/09/428,082B
;  CURRENT FILING DATE:  1999-10-22
;  PRIOR APPLICATION NUMBER: 60/105,371
;  PRIOR FILING DATE: 1998-10-23
;  NUMBER OF SEQ ID NOS: 1133
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
;   LENGTH: 277
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: EMP-EMP-Fc
US-09-428-082B-20

  Query Match             79.5%;  Score 97;  DB 2;  Length 277;
  Best Local Similarity   42.1%;  Pred. No. 1.6e-06;
  Matches   16;  Conservative    0;  Mismatches   12;  Indels   10;  Gaps    1;

Qy          1 YXCXXGPXTWXCXP----------YXCXXGPXTWXCXP 28
              | |  || || | |           | |  || || | |
Db          5 YSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKP 42


RESULT 8
US-09-428-082B-22
; Sequence 22, Application US/09428082B
; Patent No. 6660843
; GENERAL INFORMATION:
;  APPLICANT: FEIGE, ULRICH
;  APPLICANT:  LIU, CHUAN-FA
;  APPLICANT:  CHEETHAM, JANET C.
;  APPLICANT:  BOONE, THOMAS CHARLES
;  TITLE OF INVENTION: MODIFIED PEPTIDES AS THERAPEUTIC AGENTS
;  FILE REFERENCE: A-527
;  CURRENT APPLICATION NUMBER: US/09/428,082B
;  CURRENT FILING DATE:  1999-10-22
;  PRIOR APPLICATION NUMBER: 60/105,371
;  PRIOR FILING DATE: 1998-10-23
;  NUMBER OF SEQ ID NOS: 1133
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22
;   LENGTH: 277
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Fc-EMP-EMP
US-09-428-082B-22

  Query Match             79.5%;  Score 97;  DB 2;  Length 277;
  Best Local Similarity   42.1%;  Pred. No. 1.6e-06;
  Matches   16;  Conservative    0;  Mismatches   12;  Indels   10;  Gaps    1;

Qy          1 YXCXXGPXTWXCXP----------YXCXXGPXTWXCXP 28
              | |  || || | |           | |  || || | |
Db        237 YSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKP 274


RESULT 9
US-08-484-135-24
; Sequence 24, Application US/08484135
; Patent No. 5767078
;  GENERAL INFORMATION:
;    APPLICANT:  Johnson, Dana L
;    APPLICANT:  Zivin, Robert A
;    TITLE OF INVENTION:  AGONIST PEPTIDE DIMERS
;    NUMBER OF SEQUENCES:  93
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Frank S. DiGiglio
;      STREET:  400 Garden City Plaza
;      CITY:  Garden City
;      STATE:  New York
;      COUNTRY:  U.S.A..
;      ZIP:  11530
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/484,135
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DiGiglio, Frank S
;      REGISTRATION NUMBER:  31,346
;      REFERENCE/DOCKET NUMBER:  9594
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (516) 742-4343
;      TELEFAX:  (516) 742-4366
;  INFORMATION FOR SEQ ID NO:  24:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-484-135-24

  Query Match             51.6%;  Score 63;  DB 1;  Length 20;
  Best Local Similarity   60.0%;  Pred. No. 0.0093;
  Matches    9;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 YXCXXGPXTWXCXPY 15
              | |  || || | ||
Db          4 YSCRMGPMTWVCIPY 18


RESULT 10
US-08-484-135-78
; Sequence 78, Application US/08484135
; Patent No. 5767078
;  GENERAL INFORMATION:
;    APPLICANT:  Johnson, Dana L
;    APPLICANT:  Zivin, Robert A
;    TITLE OF INVENTION:  AGONIST PEPTIDE DIMERS
```

```
;    NUMBER OF SEQUENCES:  93
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Frank S. DiGiglio
;      STREET:  400 Garden City Plaza
;      CITY:  Garden City
;      STATE:  New York
;      COUNTRY:  U.S.A..
;      ZIP:  11530
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/484,135
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DiGiglio, Frank S
;      REGISTRATION NUMBER:  31,346
;      REFERENCE/DOCKET NUMBER:  9594
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (516) 742-4343
;      TELEFAX:  (516) 742-4366
;  INFORMATION FOR SEQ ID NO:  78:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-484-135-78

  Query Match             51.6%;  Score 63;  DB 1;  Length 20;
  Best Local Similarity   60.0%;  Pred. No. 0.0093;
  Matches    9;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 YXCXXGPXTWXCXPY 15
              | |  || || | ||
Db          4 YVCRMGPMTWVCAPY 18


RESULT 11
US-08-484-635-40
; Sequence 40, Application US/08484635
; Patent No. 5773569
;  GENERAL INFORMATION:
;    APPLICANT:  Wrighton, Nicholas C.
;    APPLICANT:  Dower, William J.
;    APPLICANT:  Chang, Ray S.
;    APPLICANT:  Kashyap, Arun K.
;    APPLICANT:  Jolliffe, Linda K.
;    APPLICANT:  Johnson, Dana
;    APPLICANT:  Mulcahy, Linda
;    TITLE OF INVENTION:  Compounds and Peptides That Bind to the
;    TITLE OF INVENTION:  Erythropoietin Receptor
;    NUMBER OF SEQUENCES:  259
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew
;      STREET:  One Market Plaza, Steuart Street Tower
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94105-1492
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/484,635
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/155,940
;      FILING DATE:  19-NOV-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Garrett-Wackowski, Eugenia
;      REGISTRATION NUMBER:  37,330
;      REFERENCE/DOCKET NUMBER:  16528A-43-1-1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 543-9600
;      TELEFAX:  (415) 543-5043
;  INFORMATION FOR SEQ ID NO:  40:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-484-635-40

  Query Match             51.6%;  Score 63;  DB 1;  Length 20;
  Best Local Similarity   60.0%;  Pred. No. 0.0093;
  Matches    9;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 YXCXXGPXTWXCXPY 15
              | |  || || | ||
Db          4 YVCRMGPMTWVCAPY 18


RESULT 12
US-08-484-635-206
; Sequence 206, Application US/08484635
; Patent No. 5773569
;  GENERAL INFORMATION:
;    APPLICANT:  Wrighton, Nicholas C.
;    APPLICANT:  Dower, William J.
;    APPLICANT:  Chang, Ray S.
;    APPLICANT:  Kashyap, Arun K.
;    APPLICANT:  Jolliffe, Linda K.
;    APPLICANT:  Johnson, Dana
;    APPLICANT:  Mulcahy, Linda
;    TITLE OF INVENTION:  Compounds and Peptides That Bind to the
;    TITLE OF INVENTION:  Erythropoietin Receptor
;    NUMBER OF SEQUENCES:  259
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew
;      STREET:  One Market Plaza, Steuart Street Tower
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94105-1492
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/484,635
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/155,940
;      FILING DATE:  19-NOV-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Garrett-Wackowski, Eugenia
;      REGISTRATION NUMBER:  37,330
;      REFERENCE/DOCKET NUMBER:  16528A-43-1-1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 543-9600
;      TELEFAX:  (415) 543-5043
;  INFORMATION FOR SEQ ID NO:  206:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  amino acid
```

```
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-484-635-206

  Query Match             51.6%;  Score 63;  DB 1;  Length 20;
  Best Local Similarity   60.0%;  Pred. No. 0.0093;
  Matches    9;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 YXCXXGPXTWXCXPY 15
              | |  || || | ||
Db          4 YSCRMGPMTWVCIPY 18


RESULT 13
US-08-484-631-40
; Sequence 40, Application US/08484631
; Patent No. 5830851
;  GENERAL INFORMATION:
;    APPLICANT:  Wrighton, Nicholas C.
;    APPLICANT:  Dower, William J.
;    APPLICANT:  Chang, Ray S.
;    APPLICANT:  Kashyap, Arun K.
;    APPLICANT:  Jolliffe, Linda K.
;    APPLICANT:  Johnson, Dana
;    APPLICANT:  Mulcahy, Linda
;    TITLE OF INVENTION:  Compounds and Peptides That Bind to the
;    TITLE OF INVENTION:  Erythropoietin Receptor
;    NUMBER OF SEQUENCES:  259
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew
;      STREET:  One Market Plaza, Steuart Street Tower
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94105-1492
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/484,631
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/155,940
;      FILING DATE:  19-NOV-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Garrett-Wackowski, Eugenia
;      REGISTRATION NUMBER:  37,330
;      REFERENCE/DOCKET NUMBER:  16528A-43-1-2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 543-9600
;      TELEFAX:  (415) 543-5043
;  INFORMATION FOR SEQ ID NO:  40:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-484-631-40

  Query Match             51.6%;  Score 63;  DB 1;  Length 20;
  Best Local Similarity   60.0%;  Pred. No. 0.0093;
  Matches    9;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 YXCXXGPXTWXCXPY 15
              | |  || || | ||
Db          4 YVCRMGPMTWVCAPY 18


RESULT 14
US-08-484-631-206
; Sequence 206, Application US/08484631
; Patent No. 5830851
;  GENERAL INFORMATION:
;    APPLICANT:  Wrighton, Nicholas C.
;    APPLICANT:  Dower, William J.
;    APPLICANT:  Chang, Ray S.
;    APPLICANT:  Kashyap, Arun K.
;    APPLICANT:  Jolliffe, Linda K.
;    APPLICANT:  Johnson, Dana
;    APPLICANT:  Mulcahy, Linda
;    TITLE OF INVENTION:  Compounds and Peptides That Bind to the
;    TITLE OF INVENTION:  Erythropoietin Receptor
;    NUMBER OF SEQUENCES:  259
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew
;      STREET:  One Market Plaza, Steuart Street Tower
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94105-1492
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/484,631
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/155,940
;      FILING DATE:  19-NOV-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Garrett-Wackowski, Eugenia
;      REGISTRATION NUMBER:  37,330
;      REFERENCE/DOCKET NUMBER:  16528A-43-1-2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 543-9600
;      TELEFAX:  (415) 543-5043
;  INFORMATION FOR SEQ ID NO:  206:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-484-631-206

  Query Match             51.6%;  Score 63;  DB 1;  Length 20;
  Best Local Similarity   60.0%;  Pred. No. 0.0093;
  Matches    9;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 YXCXXGPXTWXCXPY 15
              | |  || || | ||
Db          4 YSCRMGPMTWVCIPY 18


RESULT 15
US-08-827-570-40
; Sequence 40, Application US/08827570
; Patent No. 5986047
;  GENERAL INFORMATION:
;    APPLICANT:  Wrighton, Nicholas C.
;    APPLICANT:  Dower, William J.
;    APPLICANT:  Chang, Ray S.
;    APPLICANT:  Kashyap, Arun K.
;    APPLICANT:  Jolliffe, Linda K.
;    APPLICANT:  Johnson, Dana
;    APPLICANT:  Mulcahy, Linda
;    TITLE OF INVENTION:  Compounds and Peptides That Bind to the
;    TITLE OF INVENTION:  Erythropoietin Receptor
```

```
;    NUMBER OF SEQUENCES:  259
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew
;      STREET:  One Market Plaza, Steuart Street Tower
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94105-1492
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/827,570
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/484,635
;      FILING DATE:  07-JUN-1995
;      APPLICATION NUMBER:  US 08/155,940
;      FILING DATE:  19-NOV-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Garrett-Wackowski, Eugenia
;      REGISTRATION NUMBER:  37,330
;      REFERENCE/DOCKET NUMBER:  16528A-43-1-1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 543-9600
;      TELEFAX:  (415) 543-5043
;  INFORMATION FOR SEQ ID NO:  40:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-827-570-40

  Query Match             51.6%;  Score 63;  DB 1;  Length 20;
  Best Local Similarity   60.0%;  Pred. No. 0.0093;
  Matches    9;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 YXCXXGPXTWXCXPY 15
              | |  || || | ||
Db          4 YVCRMGPMTWVCAPY 18


Search completed: March 31, 2006, 16:40:35
Job time : 27.4428 secs
```

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: March 31, 2006, 16:09:06 ; Search time 53.9801 Seconds
(without alignments)
113.955 Million cell updates/sec

Title: US-10-609-217-85
Perfect score: 61
Sequence: 1 YXCXXGPXTWXCXP 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	58	95.1	20	2	AAY13728	Aay13728 Erythropo
2	58	95.1	20	2	AAY13687	Aay13687 Erythropo
3	58	95.1	20	2	AAW27001	Aaw27001 Monomer s
4	58	95.1	21	9	ADU91978	Adu91978 EPO-R ago
5	58	95.1	22	2	AAY13709	Aay13709 Erythropo
6	58	95.1	22	2	AAY26355	Aay26355 Erythropo
7	58	95.1	22	2	AAW27023	Aaw27023 Monomer s
8	57	93.4	17	9	ADU91963	Adu91963 EPO-R ago
9	57	93.4	17	9	ADU92005	Adu92005 EPO-R ago
10	57	93.4	20	2	AAY26409	Aay26409 Erythropo
11	57	93.4	20	2	AAY13650	Aay13650 Erythropo
12	57	93.4	20	2	AAY13676	Aay13676 Erythropo
13	57	93.4	20	2	AAY13630	Aay13630 Erythropo
14	57	93.4	20	2	AAY26383	Aay26383 Erythropo
15	57	93.4	20	2	AAW26990	Aaw26990 Monomer s
16	57	93.4	20	2	AAW26966	Aaw26966 Monomer s
17	57	93.4	20	2	AAW27042	Aaw27042 Monomer s
18	57	93.4	20	3	AAB17033	Aab17033 EPO-mimet
19	57	93.4	20	3	AAB17926	Aab17926 EPO-mimet
20	57	93.4	20	3	AAB13504	Aab13504 Erythropo
21	57	93.4	20	5	ABB72816	Abb72816 Erythropo
22	57	93.4	20	5	ABB72843	Abb72843 Erythropo
23	57	93.4	20	5	AAU74459	Aau74459 Human ery
24	57	93.4	20	5	AAU74480	Aau74480 Human ery
25	57	93.4	20	7	ADJ72555	Adj72555 EPO mimet
26	57	93.4	20	7	ADJ72582	Adj72582 EPO mimet
27	57	93.4	20	8	ADH10378	Adh10378 Erythropo
28	57	93.4	20	8	ADJ52191	Adj52191 CH1 delet
29	57	93.4	20	8	ADJ52218	Adj52218 CH1 delet
30	57	93.4	20	8	ADJ51180	Adj51180 CH1 delet
31	57	93.4	20	8	ADJ51153	Adj51153 CH1 delet
32	57	93.4	20	8	ADU47106	Adu47106 Erythropo
33	57	93.4	20	9	ADZ44405	Adz44405 Erythropo
34	57	93.4	20	9	ADZ44426	Adz44426 Erythropo
35	57	93.4	21	9	ADU91939	Adu91939 EPO-R ago
36	57	93.4	21	9	ADU91989	Adu91989 EPO-R ago
37	57	93.4	21	9	ADU91929	Adu91929 EPO-R ago
38	57	93.4	22	2	AAY13715	Aay13715 Erythropo
39	57	93.4	22	2	AAY26363	Aay26363 Erythropo
40	57	93.4	22	2	AAW27029	Aaw27029 Monomer s
41	57	93.4	26	2	AAY26445	Aay26445 Erythropo
42	57	93.4	26	2	AAY26441	Aay26441 Erythropo
43	57	93.4	133	7	ADJ73535	Adj73535 Erythropo
44	56	91.8	14	2	AAY13654	Aay13654 Erythropo
45	56	91.8	14	5	AAU74472	Aau74472 Human ery

ALIGNMENTS

RESULT 1
AAY13728
ID AAY13728 standard; peptide; 20 AA.
XX
AC AAY13728;
XX
DT 06-SEP-1999 (first entry)
XX
DE Erythropoietin receptor (EPO-R) binding peptide.
XX
KW Erythropoietin; EPO; receptor; EPO deficiency; renal failure; AIDS;
KW dialysis; anaemia; autoimmune disease; chronic inflammatory disease;
KW malignancy; beta-thalassaemia; cystic fibrosis; prematurity; blood loss;
KW spinal cord injury; aging; neoplastic disease; erythropoiesis; EPO-R.
XX
OS Synthetic.
XX
PN WO9640749-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US009810.
XX
PR 07-JUN-1995; 95US-00484631.
PR 07-JUN-1995; 95US-00484635.
XX
PA (JOHJ) JOHNSON & JOHNSON CORP.
PA (AFFY-) AFFYMAX TECHNOLOGIES NV.
XX
PI Wrighton NC, Dower WJ, Chang RS, Kashyap AK, Jolliffe LK;
PI Johnson D, Mulcahy L;
XX
DR WPI; 1997-052225/05.
XX
PT Erythropoietin receptor binding peptide - useful for treating disorders
PT characterised by deficiency of EPO, or low or defective red blood cell
PT population.
XX
PS Disclosure; Fig 2; 95pp; English.
XX
CC The invention describes a peptide of 10-40 amino acid residues which
CC binds to erythropoietin (EPO) receptor and which includes the amino acid
CC sequence Cys-Xaa1-Xaa2-Gly-Pro-Xaa3-Thr-Trp-Xaa4-Cys, where Xaa1 = Arg,
CC His, Leu or Trp, Xaa2 = Met, Phe or Ile, Xaa3 = 1 of the 20 genetically
CC coded L-amino acids, and Xaa4 = Asp, Glu, Ile, Leu or Val. Optionally,
CC the peptide may be cyclised or dimerised. The peptide can be used to
CC treat a patient having a disorder characterised by a deficiency of EPO or

CC a low or defective red blood cell population. It can be used to treat end
CC stage renal failure or dialysis; anaemia associated with AIDS, autoimmune
CC disease, chronic inflammatory diseases or malignancy; beta-thalassaemia;
CC cystic fibrosis; early anaemia of prematurity; spinal cord injury; acute
CC blood loss; aging; and neoplastic disease states accompanied by abnormal
CC erythropoiesis. The peptides can also be used as reagents for detecting
CC EPO receptors on living cells, in biological fluids, in tissue
CC homogenates, etc. Sequences AAY13662-735 are representative peptides of
CC the invention
XX
SQ Sequence 20 AA;

Query Match 95.1%; Score 58; DB 2; Length 20;
Best Local Similarity 57.1%; Pred. No. 0.015;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 YXCXXGPXTWXCXP 14
| | || || | |
Db 4 YACRMGPTTWVCSP 17

RESULT 2
AAY13687
ID AAY13687 standard; peptide; 20 AA.
XX
AC AAY13687;
XX
DT 06-SEP-1999 (first entry)
XX
DE Erythropoietin receptor (EPO-R) binding peptide.
XX
KW Erythropoietin; EPO; receptor; EPO deficiency; renal failure; AIDS;
KW dialysis; anaemia; autoimmune disease; chronic inflammatory disease;
KW malignancy; beta-thalassaemia; cystic fibrosis; prematurity; blood loss;
KW spinal cord injury; aging; neoplastic disease; erythropoiesis; EPO-R.
XX
OS Synthetic.
XX
PN WO9640749-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US009810.
XX
PR 07-JUN-1995; 95US-00484631.
PR 07-JUN-1995; 95US-00484635.
XX
PA (JOHJ) JOHNSON & JOHNSON CORP.
PA (AFFY-) AFFYMAX TECHNOLOGIES NV.
XX
PI Wrighton NC, Dower WJ, Chang RS, Kashyap AK, Jolliffe LK;
PI Johnson D, Mulcahy L;
XX
DR WPI; 1997-052225/05.
XX
PT Erythropoietin receptor binding peptide - useful for treating disorders
PT characterised by deficiency of EPO, or low or defective red blood cell
PT population.
XX
PS Disclosure; Fig 2; 95pp; English.
XX
CC The invention describes a peptide of 10-40 amino acid residues which
CC binds to erythropoietin (EPO) receptor and which includes the amino acid
CC sequence Cys-Xaa1-Xaa2-Gly-Pro-Xaa3-Thr-Trp-Xaa4-Cys, where Xaa1 = Arg,
CC His, Leu or Trp, Xaa2 = Met, Phe or Ile, Xaa3 = 1 of the 20 genetically
CC coded L-amino acids, and Xaa4 = Asp, Glu, Ile, Leu or Val. Optionally,
CC the peptide may be cyclised or dimerised. The peptide can be used to
CC treat a patient having a disorder characterised by a deficiency of EPO or
CC a low or defective red blood cell population. It can be used to treat end
CC stage renal failure or dialysis; anaemia associated with AIDS, autoimmune
CC disease, chronic inflammatory diseases or malignancy; beta-thalassaemia;
CC cystic fibrosis; early anaemia of prematurity; spinal cord injury; acute
CC blood loss; aging; and neoplastic disease states accompanied by abnormal
CC erythropoiesis. The peptides can also be used as reagents for detecting
CC EPO receptors on living cells, in biological fluids, in tissue
CC homogenates, etc. Sequences AAY13662-735 are representative peptides of
CC the invention
XX
SQ Sequence 20 AA;

Query Match 95.1%; Score 58; DB 2; Length 20;
Best Local Similarity 57.1%; Pred. No. 0.015;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 YXCXXGPXTWXCXP 14
| | || || | |
Db 4 YSCRMGPTTWVCTP 17

RESULT 3
AAW27001
ID AAW27001 standard; peptide; 20 AA.
XX
AC AAW27001;
XX
DT 11-NOV-1997 (first entry)
XX
DE Monomer subunit of erythropoietin receptor binding dimer.
XX
KW Monomer; erythropoietin; EPO; receptor; binding; dimer; activation;
KW treatment; disorder; deficiency; low; defective; red blood cell;
KW erythrocyte; population; cell surface; agonist; end stage; renal;
KW failure; dialysis; anaemia; anemia; AIDS; chronic; inflammatory; disease;
KW rheumatoid arthritis; bowel inflammation; autoimmune; transfusion.
XX
OS Synthetic.
XX
PN WO9640772-A2.
XX
PD 19-DEC-1996.
XX
PF 06-JUN-1996; 96WO-US009469.
XX
PR 07-JUN-1995; 95US-00484135.
XX
PA (JOHJ) JOHNSON & JOHNSON.
XX
PI Johnson DL, Zivin RA;
XX
DR WPI; 1997-099920/09.
XX
PT Activating cell surface receptors using peptide dimer agonists - also,
PT new dimers of erythropoietin receptor binding peptide(s) useful for
PT treating patient having disorder characterised by EPO deficiency.
XX
PS Disclosure; Fig 9; 110pp; English.
XX
CC The present peptide is a specific example of a claimed generic monomer
CC subunit of an erythropoietin (EPO) receptor binding dimer, which
CC comprises 2 EPO receptor binding monomers of 10 to 40 amino acids, and
CC activates or improves the bioactivity of the EPO cell surface receptor.
CC The dimer can be used to treat disorders resulting from EPO deficiency by
CC improving the activity of its cell surface receptor, e.g. end stage renal
CC failure/dialysis, anaemia associated with AIDS or chronic inflammatory
CC diseases such as rheumatoid arthritis and chronic bowel inflammation and
CC autoimmune disease. It can also be used to boost the red cell count of a
CC patient prior to surgery or as pretreatment to transfusion. The dimer
CC peptide exhibits increased biological potency in vitro and in vivo
CC relative to its component monomeric agonists. Dimerisation may also
CC convert cell surface receptor antagonists into agonists
XX
SQ Sequence 20 AA;

Query Match 95.1%; Score 58; DB 2; Length 20;
Best Local Similarity 57.1%; Pred. No. 0.015;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
Qy          1 YXCXXGPXTWXCXP 14
              | |  || || | |
Db          4 YSCRMGPTTWVCTP 17


RESULT 4
ADU91978
ID   ADU91978 standard; peptide; 21 AA.
XX
AC   ADU91978;
XX
DT   10-FEB-2005  (first entry)
XX
DE   EPO-R agonist SEQ ID NO 119.
XX
KW   erythropoietin receptor; EPO-R; erythropoietin; renal failure;
KW   autoimmune disease; cystic fibrosis; anemia; inflammation;
KW   spinal cord injury; aging; neurological disease; nephrotropic;
KW   antianemic; immunosuppressive; CNS-Gen.; neuroprotective;
KW   respiratory-Gen.; antiinflammatory; vulnerary; nootropic; cytostatic;
KW   hemostatic; cyclic.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "Acetylated residue"
FT   Disulfide-bond  7. .16
FT   Modified-site   21
FT                   /note= "C-terminal amide"
XX
PN   WO2004101611-A2.
XX
PD   25-NOV-2004.
XX
PF   12-MAY-2004; 2004WO-US014886.
XX
PR   12-MAY-2003; 2003US-0470245P.
XX
PA   (AFFY-) AFFYMAX INC.
XX
PI   Yin K,  Holmes C,  Lalonde G,  Balu P,  Schatz PJ,  Tumelty D;
XX
DR   WPI; 2005-039329/04.
XX
PT   New peptide comprising specified sequence of amino acid is erythropoietin
PT   receptor agonist useful for treating e.g. anemia, beta-thalassemia, renal
PT   disorders.
XX
PS   Disclosure; SEQ ID NO 119; 83pp; English.
XX
CC   This invention describes a novel peptide which is an erythropoietin
CC   receptor (EPO-R) activator. The peptide forms a dimer comprising a
CC   linking moiety connecting two peptide chains composed of ADU91861. The N-
CC   terminal of the peptide is acetylated. The EPO-R activator further
CC   comprises at least one water soluble polymer, preferably polyethylene
CC   glycol (PEG) covalently bound to the peptide and a spacer moiety. The
CC   products of the invention are used for treating disorders associated with
CC   deficiency of erythropoietin or low or defective red blood cell
CC   population, end stage renal failure or dialysis, anemia associated with
CC   AIDS, autoimmune disease or malignancy, beta-thalassemia, cystic
CC   fibrosis, early anemia of prematurity, anemia associated with chronic
CC   inflammatory disease, spinal cord injury, acute blood loss, aging and
CC   neoplastic disease states accompanied by abnormal erythropoiesis. The
CC   peptide compounds are potent agonists of erythropoietin receptor and have
CC   nephrotropic, antianemic, immunosuppressive, CNS-Gen., neuroprotective,
CC   respiratory-Gen., antiinflammatory, vulnerary, nootropic, cytostatic and
CC   hemostatic activity. This sequence represents a peptide which acts as an
CC   erythropoietin receptor (EPO-R) agonist.
XX
SQ   Sequence 21 AA;

  Query Match             95.1%;  Score 58;  DB 9;  Length 21;
  Best Local Similarity   57.1%;  Pred. No. 0.016;
  Matches    8;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 YXCXXGPXTWXCXP 14
              | |  || || | |
Db          5 YSCQMGPTTWTCVP 18


RESULT 5
AAY13709
ID   AAY13709 standard; peptide; 22 AA.
XX
AC   AAY13709;
XX
DT   06-SEP-1999  (first entry)
XX
DE   Erythropoietin receptor (EPO-R) binding peptide.
XX
KW   Erythropoietin; EPO; receptor; EPO deficiency; renal failure; AIDS;
KW   dialysis; anaemia; autoimmune disease; chronic inflammatory disease;
KW   malignancy; beta-thalassaemia; cystic fibrosis; prematurity; blood loss;
KW   spinal cord injury; aging; neoplastic disease; erythropoiesis; EPO-R.
XX
OS   Synthetic.
XX
PN   WO9640749-A1.
XX
PD   19-DEC-1996.
XX
PF   07-JUN-1996;   96WO-US009810.
XX
PR   07-JUN-1995;   95US-00484631.
PR   07-JUN-1995;   95US-00484635.
XX
PA   (JOHJ ) JOHNSON & JOHNSON CORP.
PA   (AFFY-) AFFYMAX TECHNOLOGIES NV.
XX
PI   Wrighton NC,  Dower WJ,  Chang RS,  Kashyap AK,  Jolliffe LK;
PI   Johnson D,  Mulcahy L;
XX
DR   WPI; 1997-052225/05.
XX
PT   Erythropoietin receptor binding peptide - useful for treating disorders
PT   characterised by deficiency of EPO, or low or defective red blood cell
PT   population.
XX
PS   Disclosure; Fig 2; 95pp; English.
XX
CC   The invention describes a peptide of 10-40 amino acid residues which
CC   binds to erythropoietin (EPO) receptor and which includes the amino acid
CC   sequence Cys-Xaa1-Xaa2-Gly-Pro-Xaa3-Thr-Trp-Xaa4-Cys, where Xaa1 = Arg,
CC   His, Leu or Trp, Xaa2 = Met, Phe or Ile, Xaa3 = 1 of the 20 genetically
CC   coded L-amino acids, and Xaa4 = Asp, Glu, Ile, Leu or Val. Optionally,
CC   the peptide may be cyclised or dimerised. The peptide can be used to
CC   treat a patient having a disorder characterised by a deficiency of EPO or
CC   a low or defective red blood cell population. It can be used to treat end
CC   stage renal failure or dialysis; anaemia associated with AIDS, autoimmune
CC   disease, chronic inflammatory diseases or malignancy; beta-thalassaemia;
CC   cystic fibrosis; early anaemia of prematurity; spinal cord injury; acute
CC   blood loss; aging; and neoplastic disease states accompanied by abnormal
CC   erythropoiesis. The peptides can also be used as reagents for detecting
CC   EPO receptors on living cells, in biological fluids, in tissue
CC   homogenates, etc. Sequences AAY13662-735 are representative peptides of
CC   the invention
XX
SQ   Sequence 22 AA;

  Query Match             95.1%;  Score 58;  DB 2;  Length 22;
  Best Local Similarity   57.1%;  Pred. No. 0.017;
  Matches    8;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 YXCXXGPXTWXCXP 14
```

```
              | |  || || | |
Db          4 YSCFMGPTTWVCSP 17


RESULT 6
AAY26355
ID   AAY26355 standard; peptide; 22 AA.
XX
AC   AAY26355;
XX
DT   06-SEP-1999  (first entry)
XX
DE   Erythropoietin receptor (EPO-R) binding peptide.
XX
KW   Erythropoietin; EPO; receptor; EPO deficiency; renal failure; AIDS;
KW   dialysis; anaemia; autoimmune disease; chronic inflammatory disease;
KW   malignancy; beta-thalassaemia; cystic fibrosis; prematurity; blood loss;
KW   spinal cord injury; aging; neoplastic disease; erythropoiesis; EPO-R.
XX
OS   Synthetic.
XX
PN   WO9640749-A1.
XX
PD   19-DEC-1996.
XX
PF   07-JUN-1996;   96WO-US009810.
XX
PR   07-JUN-1995;   95US-00484631.
PR   07-JUN-1995;   95US-00484635.
XX
PA   (JOHJ ) JOHNSON & JOHNSON CORP.
PA   (AFFY-) AFFYMAX TECHNOLOGIES NV.
XX
PI   Wrighton NC,  Dower WJ,  Chang RS,  Kashyap AK,  Jolliffe LK;
PI   Johnson D,  Mulcahy L;
XX
DR   WPI; 1997-052225/05.
XX
PT   Erythropoietin receptor binding peptide - useful for treating disorders
PT   characterised by deficiency of EPO, or low or defective red blood cell
PT   population.
XX
PS   Disclosure; Page 16; 95pp; English.
XX
CC   The invention describes a peptide of 10-40 amino acid residues which
CC   binds to erythropoietin (EPO) receptor and which includes the amino acid
CC   sequence Cys-Xaa1-Xaa2-Gly-Pro-Xaa3-Thr-Trp-Xaa4-Cys, where Xaa1 = Arg,
CC   His, Leu or Trp, Xaa2 = Met, Phe or Ile, Xaa3 = 1 of the 20 genetically
CC   coded L-amino acids, and Xaa4 = Asp, Glu, Ile, Leu or Val. Optionally,
CC   the peptide may be cyclised or dimerised. The peptide can be used to
CC   treat a patient having a disorder characterised by a deficiency of EPO or
CC   a low or defective red blood cell population. It can be used to treat end
CC   stage renal failure or dialysis; anaemia associated with AIDS, autoimmune
CC   disease, chronic inflammatory diseases or malignancy; beta-thalassaemia;
CC   cystic fibrosis; early anaemia of prematurity; spinal cord injury; acute
CC   blood loss; aging; and neoplastic disease states accompanied by abnormal
CC   erythropoiesis. The peptides can also be used as reagents for detecting
CC   EPO receptors on living cells, in biological fluids, in tissue
CC   homogenates, etc. Sequences AAY26352-548 are representative peptides
CC   falling within the above peptide motif and isolated during the affinity
CC   selection process
XX
SQ   Sequence 22 AA;

  Query Match             95.1%;  Score 58;  DB 2;  Length 22;
  Best Local Similarity   57.1%;  Pred. No. 0.017;
  Matches    8;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 YXCXXGPXTWXCXP 14
              | |  || || | |
Db          4 YSCFMGPTTWVCSP 17


RESULT 7
AAW27023
ID   AAW27023 standard; peptide; 22 AA.
XX
AC   AAW27023;
XX
DT   11-NOV-1997  (first entry)
XX
DE   Monomer subunit of erythropoietin receptor binding dimer.
XX
KW   Monomer; erythropoietin; EPO; receptor; binding; dimer; activation;
KW   treatment; disorder; deficiency; low; defective; red blood cell;
KW   erythrocyte; population; cell surface; agonist; end stage; renal;
KW   failure; dialysis; anaemia; anemia; AIDS; chronic; inflammatory; disease;
KW   rheumatoid arthritis; bowel inflammation; autoimmune; transfusion.
XX
OS   Synthetic.
XX
PN   WO9640772-A2.
XX
PD   19-DEC-1996.
XX
PF   06-JUN-1996;   96WO-US009469.
XX
PR   07-JUN-1995;   95US-00484135.
XX
PA   (JOHJ ) JOHNSON & JOHNSON.
XX
PI   Johnson DL,  Zivin RA;
XX
DR   WPI; 1997-099920/09.
XX
PT   Activating cell surface receptors using peptide dimer agonists - also,
PT   new dimers of erythropoietin receptor binding peptide(s) useful for
PT   treating patient having disorder characterised by EPO deficiency.
XX
PS   Disclosure; Fig 9; 110pp; English.
XX
CC   The present peptide is a specific example of a claimed generic monomer
CC   subunit of an erythropoietin (EPO) receptor binding dimer, which
CC   comprises 2 EPO receptor binding monomers of 10 to 40 amino acids, and
CC   activates or improves the bioactivity of the EPO cell surface receptor.
CC   The dimer can be used to treat disorders resulting from EPO deficiency by
CC   improving the activity of its cell surface receptor, e.g. end stage renal
CC   failure/dialysis, anaemia associated with AIDS or chronic inflammatory
CC   diseases such as rheumatoid arthritis and chronic bowel inflammation and
CC   autoimmune disease. It can also be used to boost the red cell count of a
CC   patient prior to surgery or as pretreatment to transfusion. The dimer
CC   peptide exhibits increased biological potency in vitro and in vivo
CC   relative to its component monomeric agonists. Dimerisation may also
CC   convert cell surface receptor antagonists into agonists
XX
SQ   Sequence 22 AA;

  Query Match             95.1%;  Score 58;  DB 2;  Length 22;
  Best Local Similarity   57.1%;  Pred. No. 0.017;
  Matches    8;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 YXCXXGPXTWXCXP 14
              | |  || || | |
Db          4 YSCFMGPTTWVCSP 17


RESULT 8
ADU91963
ID   ADU91963 standard; peptide; 17 AA.
XX
AC   ADU91963;
XX
DT   10-FEB-2005  (first entry)
XX
DE   EPO-R agonist SEQ ID NO 104.
XX
```

```
KW   erythropoietin receptor; EPO-R; erythropoietin; renal failure;
KW   autoimmune disease; cystic fibrosis; anemia; inflammation;
KW   spinal cord injury; aging; neurological disease; nephrotropic;
KW   antianemic; immunosuppressive; CNS-Gen.; neuroprotective;
KW   respiratory-Gen.; antiinflammatory; vulnerary; nootropic; cytostatic;
KW   hemostatic; cyclic.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "Acetylated residue"
FT   Disulfide-bond  4. .13
FT   Modified-site   17
FT                   /note= "C-terminal amide"
XX
PN   WO2004101611-A2.
XX
PD   25-NOV-2004.
XX
PF   12-MAY-2004; 2004WO-US014886.
XX
PR   12-MAY-2003; 2003US-0470245P.
XX
PA   (AFFY-) AFFYMAX INC.
XX
PI   Yin K,  Holmes C,  Lalonde G,  Balu P,  Schatz PJ,  Tumelty D;
XX
DR   WPI; 2005-039329/04.
XX
PT   New peptide comprising specified sequence of amino acid is erythropoietin
PT   receptor agonist useful for treating e.g. anemia, beta-thalassemia, renal
PT   disorders.
XX
PS   Disclosure; SEQ ID NO 104; 83pp; English.
XX
CC   This invention describes a novel peptide which is an erythropoietin
CC   receptor (EPO-R) activator. The peptide forms a dimer comprising a
CC   linking moiety connecting two peptide chains composed of ADU91861. The N-
CC   terminal of the peptide is acetylated. The EPO-R activator further
CC   comprises at least one water soluble polymer, preferably polyethylene
CC   glycol (PEG) covalently bound to the peptide and a spacer moiety. The
CC   products of the invention are used for treating disorders associated with
CC   deficiency of erythropoietin or low or defective red blood cell
CC   population, end stage renal failure or dialysis, anemia associated with
CC   AIDS, autoimmune disease or malignancy, beta-thalassemia, cystic
CC   fibrosis, early anemia of prematurity, anemia associated with chronic
CC   inflammatory disease, spinal cord injury, acute blood loss, aging and
CC   neoplastic disease states accompanied by abnormal erythropoiesis. The
CC   peptide compounds are potent agonists of erythropoietin receptor and have
CC   nephrotropic, antianemic, immunosuppressive, CNS-Gen., neuroprotective,
CC   respiratory-Gen., antiinflammatory, vulnerary, nootropic, cytostatic and
CC   hemostatic activity. This sequence represents a peptide which acts as an
CC   erythropoietin receptor (EPO-R) agonist.
XX
SQ   Sequence 17 AA;

  Query Match             93.4%;  Score 57;  DB 9;  Length 17;
  Best Local Similarity   57.1%;  Pred. No. 0.019;
  Matches    8;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 YXCXXGPXTWXCXP 14
              | |  || || | |
Db          2 YSCRMGPMTWVCSP 15


RESULT 9
ADU92005
ID   ADU92005 standard; peptide; 17 AA.
XX
AC   ADU92005;
XX
DT   10-FEB-2005  (first entry)
XX
DE   EPO-R agonist SEQ ID NO 146.
XX
KW   erythropoietin receptor; EPO-R; erythropoietin; renal failure;
KW   autoimmune disease; cystic fibrosis; anemia; inflammation;
KW   spinal cord injury; aging; neurological disease; nephrotropic;
KW   antianemic; immunosuppressive; CNS-Gen.; neuroprotective;
KW   respiratory-Gen.; antiinflammatory; vulnerary; nootropic; cytostatic;
KW   hemostatic; cyclic.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "Acetylated residue"
FT   Disulfide-bond  4. .13
FT   Modified-site   17
FT                   /note= "C-terminal amide"
XX
PN   WO2004101611-A2.
XX
PD   25-NOV-2004.
XX
PF   12-MAY-2004; 2004WO-US014886.
XX
PR   12-MAY-2003; 2003US-0470245P.
XX
PA   (AFFY-) AFFYMAX INC.
XX
PI   Yin K,  Holmes C,  Lalonde G,  Balu P,  Schatz PJ,.  Tumelty D;
XX
DR   WPI; 2005-039329/04.
XX
PT   New peptide comprising specified sequence of amino acid is erythropoietin
PT   receptor agonist useful for treating e.g. anemia, beta-thalassemia, renal
PT   disorders.
XX
PS   Disclosure; SEQ ID NO 146; 83pp; English.
XX
CC   This invention describes a novel peptide which is an erythropoietin
CC   receptor (EPO-R) activator. The peptide forms a dimer comprising a
CC   linking moiety connecting two peptide chains composed of ADU91861. The N-
CC   terminal of the peptide is acetylated. The EPO-R activator further
CC   comprises at least one water soluble polymer, preferably polyethylene
CC   glycol (PEG) covalently bound to the peptide and a spacer moiety. The
CC   products of the invention are used for treating disorders associated with
CC   deficiency of erythropoietin or low or defective red blood cell
CC   population, end stage renal failure or dialysis, anemia associated with
CC   AIDS, autoimmune disease or malignancy, beta-thalassemia, cystic
CC   fibrosis, early anemia of prematurity, anemia associated with chronic
CC   inflammatory disease, spinal cord injury, acute blood loss, aging and
CC   neoplastic disease states accompanied by abnormal erythropoiesis. The
CC   peptide compounds are potent agonists of erythropoietin receptor and have
CC   nephrotropic, antianemic, immunosuppressive, CNS-Gen., neuroprotective,
CC   respiratory-Gen., antiinflammatory, vulnerary, nootropic, cytostatic and
CC   hemostatic activity. This sequence represents a peptide which acts as an
CC   erythropoietin receptor (EPO-R) agonist.
XX
SQ   Sequence 17 AA;

  Query Match             93.4%;  Score 57;  DB 9;  Length 17;
  Best Local Similarity   57.1%;  Pred. No. 0.019;
  Matches    8;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 YXCXXGPXTWXCXP 14
              | |  || || | |
Db          2 YTCRFGPLTWECTP 15


RESULT 10
AAY26409
ID   AAY26409 standard; peptide; 20 AA.
XX
```

```
AC   AAY26409;
XX
DT   06-SEP-1999  (first entry)
XX
DE   Erythropoietin receptor (EPO-R) binding peptide.
XX
KW   Erythropoietin; EPO; receptor; EPO deficiency; renal failure; AIDS;
KW   dialysis; anaemia; autoimmune disease; chronic inflammatory disease;
KW   malignancy; beta-thalassaemia; cystic fibrosis; prematurity; blood loss;
KW   spinal cord injury; aging; neoplastic disease; erythropoiesis; EPO-R.
XX
OS   Synthetic.
XX
PN   WO9640749-A1.
XX
PD   19-DEC-1996.
XX
PF   07-JUN-1996;   96WO-US009810.
XX
PR   07-JUN-1995;   95US-00484631.
PR   07-JUN-1995;   95US-00484635.
XX
PA   (JOHJ ) JOHNSON & JOHNSON CORP.
PA   (AFFY-) AFFYMAX TECHNOLOGIES NV.
XX
PI   Wrighton NC,  Dower WJ,  Chang RS,  Kashyap AK,  Jolliffe LK;
PI   Johnson D,  Mulcahy L;
XX
DR   WPI; 1997-052225/05.
XX
PT   Erythropoietin receptor binding peptide - useful for treating disorders
PT   characterised by deficiency of EPO, or low or defective red blood cell
PT   population.
XX
PS   Disclosure; Page 19; 95pp; English.
XX
CC   The invention describes a peptide of 10-40 amino acid residues which
CC   binds to erythropoietin (EPO) receptor and which includes the amino acid
CC   sequence Cys-Xaa1-Xaa2-Gly-Pro-Xaa3-Thr-Trp-Xaa4-Cys, where Xaa1 = Arg,
CC   His, Leu or Trp, Xaa2 = Met, Phe or Ile, Xaa3 = 1 of the 20 genetically
CC   coded L-amino acids, and Xaa4 = Asp, Glu, Ile, Leu or Val. Optionally,
CC   the peptide may be cyclised or dimerised. The peptide can be used to
CC   treat a patient having a disorder characterised by a deficiency of EPO or
CC   a low or defective red blood cell population. It can be used to treat end
CC   stage renal failure or dialysis; anaemia associated with AIDS, autoimmune
CC   disease, chronic inflammatory diseases or malignancy; beta-thalassaemia;
CC   cystic fibrosis; early anaemia of prematurity; spinal cord injury; acute
CC   blood loss; aging; and neoplastic disease states accompanied by abnormal
CC   erythropoiesis. The peptides can also be used as reagents for detecting
CC   EPO receptors on living cells, in biological fluids, in tissue
CC   homogenates, etc. Sequences AAY26352-548 are representative peptides
CC   falling within the above peptide motif and isolated during the affinity
CC   selection process
XX
SQ   Sequence 20 AA;

  Query Match             93.4%;  Score 57;  DB 2;  Length 20;
  Best Local Similarity   57.1%;  Pred. No. 0.022;
  Matches    8;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 YXCXXGPXTWXCXP 14
              | |  || || | |
Db          4 YACRMGPITWVCSP 17


RESULT 11
AAY13650
ID   AAY13650 standard; peptide; 20 AA.
XX
AC   AAY13650;
XX
DT   06-SEP-1999  (first entry)
XX
DE   Erythropoietin receptor (EPO-R) binding peptide.
XX
KW   Erythropoietin; EPO; receptor; EPO deficiency; renal failure; AIDS;
KW   dialysis; anaemia; autoimmune disease; chronic inflammatory disease;
KW   malignancy; beta-thalassaemia; cystic fibrosis; prematurity; blood loss;
KW   spinal cord injury; aging; neoplastic disease; erythropoiesis; EPO-R.
XX
OS   Synthetic.
XX
PN   WO9640749-A1.
XX
PD   19-DEC-1996.
XX
PF   07-JUN-1996;   96WO-US009810.
XX
PR   07-JUN-1995;   95US-00484631.
PR   07-JUN-1995;   95US-00484635.
XX
PA   (JOHJ ) JOHNSON & JOHNSON CORP.
PA   (AFFY-) AFFYMAX TECHNOLOGIES NV.
XX
PI   Wrighton NC,  Dower WJ,  Chang RS,  Kashyap AK,  Jolliffe LK;
PI   Johnson D,  Mulcahy L;
XX
DR   WPI; 1997-052225/05.
XX
PT   Erythropoietin receptor binding peptide - useful for treating disorders
PT   characterised by deficiency of EPO, or low or defective red blood cell
PT   population.
XX
PS   Claim 6; Page 68; 95pp; English.
XX
CC   The invention describes a peptide of 10-40 amino acid residues which
CC   binds to erythropoietin (EPO) receptor and which includes the amino acid
CC   sequence Cys-Xaa1-Xaa2-Gly-Pro-Xaa3-Thr-Trp-Xaa4-Cys, where Xaa1 = Arg,
CC   His, Leu or Trp, Xaa2 = Met, Phe or Ile, Xaa3 = 1 of the 20 genetically
CC   coded L-amino acids, and Xaa4 = Asp, Glu, Ile, Leu or Val. Optionally,
CC   the peptide may be cyclised or dimerised. The peptide can be used to
CC   treat a patient having a disorder characterised by a deficiency of EPO or
CC   a low or defective red blood cell population. It can be used to treat end
CC   stage renal failure or dialysis; anaemia associated with AIDS, autoimmune
CC   disease, chronic inflammatory diseases or malignancy; beta-thalassaemia;
CC   cystic fibrosis; early anaemia of prematurity; spinal cord injury; acute
CC   blood loss; aging; and neoplastic disease states accompanied by abnormal
CC   erythropoiesis. The peptides can also be used as reagents for detecting
CC   EPO receptors on living cells, in biological fluids, in tissue
CC   homogenates, etc. Sequences AAY13624-661 represent specifc examples of
CC   EPO-R binding peptides
XX
SQ   Sequence 20 AA;

  Query Match             93.4%;  Score 57;  DB 2;  Length 20;
  Best Local Similarity   57.1%;  Pred. No. 0.022;
  Matches    8;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 YXCXXGPXTWXCXP 14
              | |  || || | |
Db          4 YSCHFGPATWVCKP 17


RESULT 12
AAY13676
ID   AAY13676 standard; peptide; 20 AA.
XX
AC   AAY13676;
XX
DT   06-SEP-1999  (first entry)
XX
DE   Erythropoietin receptor (EPO-R) binding peptide.
XX
KW   Erythropoietin; EPO; receptor; EPO deficiency; renal failure; AIDS;
KW   dialysis; anaemia; autoimmune disease; chronic inflammatory disease;
KW   malignancy; beta-thalassaemia; cystic fibrosis; prematurity; blood loss;
```

KW spinal cord injury; aging; neoplastic disease; erythropoiesis; EPO-R.
XX
OS Synthetic.
XX
PN WO9640749-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US009810.
XX
PR 07-JUN-1995; 95US-00484631.
PR 07-JUN-1995; 95US-00484635.
XX
PA (JOHJ) JOHNSON & JOHNSON CORP.
PA (AFFY-) AFFYMAX TECHNOLOGIES NV.
XX
PI Wrighton NC, Dower WJ, Chang RS, Kashyap AK, Jolliffe LK;
PI Johnson D, Mulcahy L;
XX
DR WPI; 1997-052225/05.
XX
PT Erythropoietin receptor binding peptide - useful for treating disorders
PT characterised by deficiency of EPO, or low or defective red blood cell
PT population.
XX
PS Disclosure; Fig 2; 95pp; English.
XX
CC The invention describes a peptide of 10-40 amino acid residues which
CC binds to erythropoietin (EPO) receptor and which includes the amino acid
CC sequence Cys-Xaa1-Xaa2-Gly-Pro-Xaa3-Thr-Trp-Xaa4-Cys, where Xaa1 = Arg,
CC His, Leu or Trp, Xaa2 = Met, Phe or Ile, Xaa3 = 1 of the 20 genetically
CC coded L-amino acids, and Xaa4 = Asp, Glu, Ile, Leu or Val. Optionally,
CC the peptide may be cyclised or dimerised. The peptide can be used to
CC treat a patient having a disorder characterised by a deficiency of EPO or
CC a low or defective red blood cell population. It can be used to treat end
CC stage renal failure or dialysis; anaemia associated with AIDS, autoimmune
CC disease, chronic inflammatory diseases or malignancy; beta-thalassaemia;
CC cystic fibrosis; early anaemia of prematurity; spinal cord injury; acute
CC blood loss; aging; and neoplastic disease states accompanied by abnormal
CC erythropoiesis. The peptides can also be used as reagents for detecting
CC EPO receptors on living cells, in biological fluids, in tissue
CC homogenates, etc. Sequences AAY13662-735 are representative peptides of
CC the invention
XX
SQ Sequence 20 AA;

Query Match 93.4%; Score 57; DB 2; Length 20;
Best Local Similarity 57.1%; Pred. No. 0.022;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 YXCXXGPXTWXCXP 14
| | || || | |
Db 4 YSCRMGPMTWVCSP 17

RESULT 13
AAY13630
ID AAY13630 standard; peptide; 20 AA.
XX
AC AAY13630;
XX
DT 06-SEP-1999 (first entry)
XX
DE Erythropoietin receptor (EPO-R) binding peptide.
XX
KW Erythropoietin; EPO; receptor; EPO deficiency; renal failure; AIDS;
KW dialysis; anaemia; autoimmune disease; chronic inflammatory disease;
KW malignancy; beta-thalassaemia; cystic fibrosis; prematurity; blood loss;
KW spinal cord injury; aging; neoplastic disease; erythropoiesis; EPO-R.
XX
OS Synthetic.
XX
PN WO9640749-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US009810.
XX
PR 07-JUN-1995; 95US-00484631.
PR 07-JUN-1995; 95US-00484635.
XX
PA (JOHJ) JOHNSON & JOHNSON CORP.
PA (AFFY-) AFFYMAX TECHNOLOGIES NV.
XX
PI Wrighton NC, Dower WJ, Chang RS, Kashyap AK, Jolliffe LK;
PI Johnson D, Mulcahy L;
XX
DR WPI; 1997-052225/05.
XX
PT Erythropoietin receptor binding peptide - useful for treating disorders
PT characterised by deficiency of EPO, or low or defective red blood cell
PT population.
XX
PS Claim 6; Page 68; 95pp; English.
XX
CC The invention describes a peptide of 10-40 amino acid residues which
CC binds to erythropoietin (EPO) receptor and which includes the amino acid
CC sequence Cys-Xaa1-Xaa2-Gly-Pro-Xaa3-Thr-Trp-Xaa4-Cys, where Xaa1 = Arg,
CC His, Leu or Trp, Xaa2 = Met, Phe or Ile, Xaa3 = 1 of the 20 genetically
CC coded L-amino acids, and Xaa4 = Asp, Glu, Ile, Leu or Val. Optionally,
CC the peptide may be cyclised or dimerised. The peptide can be used to
CC treat a patient having a disorder characterised by a deficiency of EPO or
CC a low or defective red blood cell population. It can be used to treat end
CC stage renal failure or dialysis; anaemia associated with AIDS, autoimmune
CC disease, chronic inflammatory diseases or malignancy; beta-thalassaemia;
CC cystic fibrosis; early anaemia of prematurity; spinal cord injury; acute
CC blood loss; aging; and neoplastic disease states accompanied by abnormal
CC erythropoiesis. The peptides can also be used as reagents for detecting
CC EPO receptors on living cells, in biological fluids, in tissue
CC homogenates, etc. Sequences AAY13624-661 represent specifc examples of
CC EPO-R binding peptides
XX
SQ Sequence 20 AA;

Query Match 93.4%; Score 57; DB 2; Length 20;
Best Local Similarity 57.1%; Pred. No. 0.022;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 YXCXXGPXTWXCXP 14
| | || || | |
Db 4 YACRMGPITWVCSP 17

RESULT 14
AAY26383
ID AAY26383 standard; peptide; 20 AA.
XX
AC AAY26383;
XX
DT 06-SEP-1999 (first entry)
XX
DE Erythropoietin receptor (EPO-R) binding peptide.
XX
KW Erythropoietin; EPO; receptor; EPO deficiency; renal failure; AIDS;
KW dialysis; anaemia; autoimmune disease; chronic inflammatory disease;
KW malignancy; beta-thalassaemia; cystic fibrosis; prematurity; blood loss;
KW spinal cord injury; aging; neoplastic disease; erythropoiesis; EPO-R.
XX
OS Synthetic.
XX
PN WO9640749-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US009810.
XX

PR 07-JUN-1995; 95US-00484631.
PR 07-JUN-1995; 95US-00484635.
XX
PA (JOHJ) JOHNSON & JOHNSON CORP.
PA (AFFY-) AFFYMAX TECHNOLOGIES NV.
XX
PI Wrighton NC, Dower WJ, Chang RS, Kashyap AK, Jolliffe LK;
PI Johnson D, Mulcahy L;
XX
DR WPI; 1997-052225/05.
XX
PT Erythropoietin receptor binding peptide - useful for treating disorders
PT characterised by deficiency of EPO, or low or defective red blood cell
PT population.
XX
PS Disclosure; Page 17; 95pp; English.
XX
CC The invention describes a peptide of 10-40 amino acid residues which
CC binds to erythropoietin (EPO) receptor and which includes the amino acid
CC sequence Cys-Xaa1-Xaa2-Gly-Pro-Xaa3-Thr-Trp-Xaa4-Cys, where Xaa1 = Arg,
CC His, Leu or Trp, Xaa2 = Met, Phe or Ile, Xaa3 = 1 of the 20 genetically
CC coded L-amino acids, and Xaa4 = Asp, Glu, Ile, Leu or Val. Optionally,
CC the peptide may be cyclised or dimerised. The peptide can be used to
CC treat a patient having a disorder characterised by a deficiency of EPO or
CC a low or defective red blood cell population. It can be used to treat end
CC stage renal failure or dialysis; anaemia associated with AIDS, autoimmune
CC disease, chronic inflammatory diseases or malignancy; beta-thalassaemia;
CC cystic fibrosis; early anaemia of prematurity; spinal cord injury; acute
CC blood loss; aging; and neoplastic disease states accompanied by abnormal
CC erythropoiesis. The peptides can also be used as reagents for detecting
CC EPO receptors on living cells, in biological fluids, in tissue
CC homogenates, etc. Sequences AAY26352-548 are representative peptides
CC falling within the above peptide motif and isolated during the affinity
CC selection process
XX
SQ Sequence 20 AA;

Query Match 93.4%; Score 57; DB 2; Length 20;
Best Local Similarity 57.1%; Pred. No. 0.022;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 YXCXXGPXTWXCXP 14
|| || || | |
Db 4 YACRMGPITWVCSP 17

RESULT 15
AAW26990
ID AAW26990 standard; peptide; 20 AA.
XX
AC AAW26990;
XX
DT 11-NOV-1997 (first entry)
XX
DE Monomer subunit of erythropoietin receptor binding dimer.
XX
KW Monomer; erythropoietin; EPO; receptor; binding; dimer; activation;
KW treatment; disorder; deficiency; low; defective; red blood cell;
KW erythrocyte; population; cell surface; agonist; end stage; renal;
KW failure; dialysis; anaemia; anemia; AIDS; chronic; inflammatory; disease;
KW rheumatoid arthritis; bowel inflammation; autoimmune; transfusion.
XX
OS Synthetic.
XX
PN WO9640772-A2.
XX
PD 19-DEC-1996.
XX
PF 06-JUN-1996; 96WO-US009469.
XX
PR 07-JUN-1995; 95US-00484135.
XX
PA (JOHJ) JOHNSON & JOHNSON.
XX
PI Johnson DL, Zivin RA;
XX
DR WPI; 1997-099920/09.
XX
PT Activating cell surface receptors using peptide dimer agonists - also,
PT new dimers of erythropoietin receptor binding peptide(s) useful for
PT treating patient having disorder characterised by EPO deficiency.
XX
PS Disclosure; Fig 9; 110pp; English.
XX
CC The present peptide is a specific example of a claimed generic monomer
CC subunit of an erythropoietin (EPO) receptor binding dimer, which
CC comprises 2 EPO receptor binding monomers of 10 to 40 amino acids, and
CC activates or improves the bioactivity of the EPO cell surface receptor.
CC The dimer can be used to treat disorders resulting from EPO deficiency by
CC improving the activity of its cell surface receptor, e.g. end stage renal
CC failure/dialysis, anaemia associated with AIDS or chronic inflammatory
CC diseases such as rheumatoid arthritis and chronic bowel inflammation and
CC autoimmune disease. It can also be used to boost the red cell count of a
CC patient prior to surgery or as pretreatment to transfusion. The dimer
CC peptide exhibits increased biological potency in vitro and in vivo
CC relative to its component monomeric agonists. Dimerisation may also
CC convert cell surface receptor antagonists into agonists
XX
SQ Sequence 20 AA;

Query Match 93.4%; Score 57; DB 2; Length 20;
Best Local Similarity 57.1%; Pred. No. 0.022;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 YXCXXGPXTWXCXP 14
|| || || | |
Db 4 YSCRMGPMTWVCSP 17

Search completed: March 31, 2006, 16:22:25
Job time : 53.9801 secs

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: March 31, 2006, 16:22:51 ; Search time 8.70647 Seconds
(without alignments)
154.717 Million cell updates/sec

Title: US-10-609-217-85
Perfect score: 61
Sequence: 1 YXCXXGPXTWXCXP 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	45	73.8	19	1	EWSMAN	ancovenin - Strept
2	40	65.6	318	2	E87929	protein T22H2.6 [i
3	40	65.6	345	2	T25138	hypothetical prote
4	40	65.6	358	2	T25137	hypothetical prote
5	39	63.9	1531	1	DVHUAR	multidrug resistan
6	37	60.7	294	2	S13141	hypothetical prote
7	37	60.7	321	2	F90826	hypothetical prote
8	37	60.7	324	2	G85684	unknown protein en
9	37	60.7	415	2	PC4407	envelope protein -
10	37	60.7	460	2	S06022	regulatory protein
11	37	60.7	475	2	H84137	hypothetical prote
12	37	60.7	652	2	S25265	outer membrane pro
13	37	60.7	652	2	D82317	iron-regulated out
14	37	60.7	3175	1	RRWVEV	genome polyprotein
15	36	59.0	123	2	I52427	guanine-nucleotide
16	36	59.0	123	2	S29714	guanine-nucleotide
17	36	59.0	350	1	DEZPA	alcohol dehydrogen
18	36	59.0	466	2	A36674	transcription fact
19	36	59.0	571	1	S30253	GABA transport pro
20	35.5	58.2	4543	1	A53102	alpha-2-macroglobu
21	35	57.4	341	1	PYVZCB	spheroidin precurs
22	35	57.4	612	2	T36880	hypothetical prote
23	35	57.4	645	2	T27186	hypothetical prote
24	35	57.4	814	2	G02390	disintegrin-like m
25	35	57.4	2531	2	S18188	notch protein homo
26	35	57.4	2531	2	A46019	notch-1 protein -
27	35	57.4	2555	2	A40043	notch protein homo
28	34.5	56.6	1661	2	T31330	head-activator bin
29	34	55.7	19	1	EWSMCN	cinnamycin - Strep
30	34	55.7	78	1	EWSMYG	cinnamycin precurs
31	34	55.7	217	2	E95370	hypothetical prote
32	34	55.7	308	2	S74719	hypothetical prote
33	34	55.7	1472	2	B54774	ATP binding casset
34	34	55.7	1693	2	S76086	beta transducin-li
35	33.5	54.9	4544	1	S02392	alpha-2-macroglobu
36	33.5	54.9	4545	1	S25111	alpha-2-macroglobu
37	33	54.1	68	2	B43940	lactococcin B prec
38	33	54.1	119	2	B98236	exsI protein prote
39	33	54.1	177	2	T01705	hypothetical prote
40	33	54.1	217	2	H86188	protein T25N20.5 [
41	33	54.1	266	2	H86407	F3H9.15 protein -
42	33	54.1	292	2	G88071	protein ZK1240.5 [
43	33	54.1	326	4	S61652	hypothetical prote
44	33	54.1	410	2	S38238	hypothetical prote
45	33	54.1	449	2	AC0234	probable exported

ALIGNMENTS

RESULT 1
EWSMAN
ancovenin - Streptomyces sp. (strain A647P-2)
C;Species: Streptomyces sp.
C;Date: 12-May-1994 #sequence_revision 19-May-1994 #text_change 09-Jul-2004
C;Accession: A61284
R;Wakamiya, T.; Ueki, Y.; Shiba, T.; Kido, Y.; Motoki, Y.
Tetrahedron Lett. 26, 665-668, 1985
A;Title: The structure of ancovenin, a new peptide inhibitor of angiotensin I converting
A;Reference number: A61284
A;Accession: A61284
A;Molecule type: protein
A;Residues: 1-19 <WAK>
A;Cross-references: UNIPROT:P38655; UNIPARC:UPI0000052CC3
C;Superfamily: cinnamycin precursor
C;Keywords: antibiotic; lanthionine
F;1-18/Cross-link: (2S,3S,6R)-3-methyl-lanthionine (Cys-Thr) #status experimental
F;4-14/Cross-link: sn-(2S,6R)-lanthionine (Ser-Cys) #status experimental
F;5-11/Cross-link: (2S,3S,6R)-3-methyl-lanthionine (Cys-Thr) #status experimental
F;6/Modified site: dehydroalanine (Ser) #status experimental

Query Match 73.8%; Score 45; DB 1; Length 19;
Best Local Similarity 60.0%; Pred. No. 0.11;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy          3 CXXGPXTWXC 12
              |  ||  || |
Db          5 CSFGPLTWSC 14
```

RESULT 2
E87929
protein T22H2.6 [imported] - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Dec-2002
C;Accession: E87929
R;anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
A;Reference number: A75000; MUID:99069613; PMID:9851916
A;Note: see websites genome.wustl.edu/gsc/C_elegans/ and www_sanger.ac.uk/Projects/C_ele
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A;Accession: E87929
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-318 <STO>
A;Cross-references: UNIPARC:UPI0000177C8F; GB:chr_I; PIDN:CAB04752.1; PID:g3880056; GSPD
C;Genetics:
A;Gene: T22H2.6
A;Map position: 1
C;Superfamily: protein T22H2.6

```
  Query Match             65.6%;  Score 40;  DB 2;  Length 318;
  Best Local Similarity   50.0%;  Pred. No. 9.5;
  Matches    6;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          3 CXXGPXTWXCXP 14
              |  |  || | |
Db         71 CKLGDNTWGCCP 82


RESULT 3
T25138
hypothetical protein T22H2.6b - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T25138
R;Lennard, N.
submitted to the EMBL Data Library, November 1996
A;Reference number: Z19985
A;Accession: T25138
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-345 <WIL>
A;Cross-references: UNIPROT:Q9U362; UNIPARC:UPI000002A1D2; EMBL:Z81595; PIDN:CAB54305.1;
A;Experimental source: clone T22H2
C;Genetics:
A;Gene: CESP:T22H2.6b
A;Map position: 1
A;Introns: 93/3; 232/3; 314/3
C;Superfamily: protein T22H2.6

  Query Match             65.6%;  Score 40;  DB 2;  Length 345;
  Best Local Similarity   50.0%;  Pred. No. 10;
  Matches    6;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          3 CXXGPXTWXCXP 14
              |  |  || | |
Db        111 CKLGDNTWGCCP 122


RESULT 4
T25137
hypothetical protein T22H2.6a - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T25137
R;Lennard, N.
submitted to the EMBL Data Library, November 1996
A;Reference number: Z19985
A;Accession: T25137
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-358 <WIL>
A;Cross-references: UNIPROT:Q9U362; UNIPARC:UPI000008667D; EMBL:Z81595; PIDN:CAB54304.1;
A;Experimental source: clone T22H2
C;Genetics:
A;Gene: CESP:T22H2.6a
A;Map position: 1
A;Introns: 93/3; 232/3; 314/3
C;Superfamily: protein T22H2.6

  Query Match             65.6%;  Score 40;  DB 2;  Length 358;
  Best Local Similarity   50.0%;  Pred. No. 11;
  Matches    6;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          3 CXXGPXTWXCXP 14
              |  |  || | |
Db        111 CKLGDNTWGCCP 122


RESULT 5
DVHUAR
multidrug resistance protein (cell line H69AR) - human
N;Alternate names: multidrug resistance-associated protein (MRP)
C;Species: Homo sapiens (man)
C;Date: 30-Jun-1993 #sequence_revision 05-Dec-1998 #text_change 19-Jan-2001
C;Accession: A44231; A37495
R;Cole, S.P.C.; Bhardwaj, G.; Gerlach, J.H.; Mackie, J.E.; Grant, C.E.; Almquist, K.C.; S
Science 258, 1650-1654, 1992
A;Title: Overexpression of a transporter gene in a multidrug-resistant human lung cancer
A;Reference number: A44231; MUID:93088080; PMID:1360704
A;Accession: A44231
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 'MAPTRSGTGMSRGIPATPTSPSAFRTRSSCGCLVFTSGPV',50-1531 <CO1>
A;Cross-references: UNIPARC:UPI00001746CB; GB:L05628; NID:g1835658
A;Experimental source: small cell lung carcinoma cell line H69AR
A;Note: sequence extracted from NCBI backbone (NCBIP:119851); this sequence has been corr
R;Cole, S.P.C.; Deeley, R.G.
Science 260, 879, 1993
A;Title: Multidrug resistance-associated protein: sequence correction.
A;Reference number: A37495; MUID:93262415; PMID:8098549
A;Accession: A37495
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-60 <CO2>
A;Cross-references: UNIPARC:UPI00001746CC; GB:L05628; NID:g1835658
A;Note: sequence extracted from NCBI backbone (NCBIP:131929)
C;Genetics:
A;Gene: GDB:MRP
A;Cross-references: GDB:136335; OMIM:158343
A;Map position: 16p13.1-16p13.1
C;Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology
C;Keywords: antibiotic resistance; ATP; duplication; nucleotide binding; P-loop; transmem
F;661-844/Domain: ATP-binding cassette homology <ABC1>
F;678-685/Region: nucleotide-binding motif A (P-loop)
F;788-792/Region: nucleotide-binding motif B
F;1310-1503/Domain: ATP-binding cassette homology <ABC2>
F;1327-1334/Region: nucleotide-binding motif A (P-loop)
F;1450-1454/Region: nucleotide-binding motif B

  Query Match             63.9%;  Score 39;  DB 1;  Length 1531;
  Best Local Similarity   42.9%;  Pred. No. 55;
  Matches    6;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          1 YXCXXGPXTWXCXP 14
              |   |  || | |
Db        544 YLSAVGTFTWVCTP 557


RESULT 6
S13141
hypothetical protein (ribosomal RNA repeat region) - Giardia lamblia
C;Species: Giardia lamblia
C;Date: 06-Dec-1996 #sequence_revision 06-Dec-1996 #text_change 05-Oct-2004
C;Accession: S13141; S10886
R;Upcroft, J.A.; Healey, A.; Mitchell, R.; Boreham, P.F.L.; Upcroft, P.
Nucleic Acids Res. 18, 7077-7081, 1990
A;Title: Antigen expression from the ribosomal DNA repeat unit of Giardia intestinalis.
A;Reference number: S13141; MUID:91088287; PMID:2263466
A;Accession: S13141
A;Molecule type: DNA
A;Residues: 1-294 <UPC>
A;Cross-references: UNIPROT:Q9XZV7; UNIPARC:UPI0000177CC5; EMBL:X52949
A;Note: the source is designated as Giardia intestinalis
A;Note: readthrough of the terminator TAG is supposed to occur between residues 241-Ala a
R;Healey, A.; Mitchell, R.; Upcroft, J.A.; Boreham, P.F.L.; Upcroft, P.
Nucleic Acids Res. 18, 4006, 1990
A;Title: Complete nucleotide sequence of the ribosomal RNA tandem repeat unit from Giardi
A;Reference number: S10886; MUID:90326542; PMID:2374731
A;Accession: S10886
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-241 <HEA>
A;Cross-references: UNIPARC:UPI0000177CC6; EMBL:X52949
A;Note: the source is designated as Giardia intestinalis
A;Note: the assignment of the coding region has been revised in reference S13141
```

C;Superfamily: Proline-rich peptide P-B

Query Match 60.7%; Score 37; DB 2; Length 294;
Best Local Similarity 62.5%; Pred. No. 29;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          7 PXTWXCXP 14
              | || | |
Db         93 PRTWACLP 100
```

RESULT 7
F90826
hypothetical protein ECs1582 [imported] - Escherichia coli (strain O157:H7, substrain RI
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: F90826
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: F90826
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-321 <HAY>
A;Cross-references: UNIPROT:Q8X356; UNIPARC:UPI00000D29F6; GB:BA000007; PIDN:BAB35005.1;
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs1582

Query Match 60.7%; Score 37; DB 2; Length 321;
Best Local Similarity 62.5%; Pred. No. 32;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          7 PXTWXCXP 14
              | || | |
Db        179 PETWLCSP 186
```

RESULT 8
G85684
unknown protein encoded by prophage CP-933C [imported] - Escherichia coli (strain O157:H
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: G85684
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: G85684
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-324 <STO>
A;Cross-references: UNIPROT:Q8X3P7; UNIPARC:UPI00000D0ED9; GB:AE005174; NID:g12514761; P
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: Z1842

Query Match 60.7%; Score 37; DB 2; Length 324;
Best Local Similarity 62.5%; Pred. No. 32;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          7 PXTWXCXP 14
              | || | |
Db        179 PETWLCSP 186
```

RESULT 9
PC4407
envelope protein - hepatitis C virus (fragment)
C;Species: hepatitis C virus
C;Date: 10-Nov-1997 #sequence_revision 23-Jan-1998 #text_change 09-Jul-2004
C;Accession: PC4407
R;Li, G.; Yao, J.; Peng, W.
Chinese J. Virol. 13, 24-32, 1997
A;Title: Sequence of genomic region of hepatitis C virus envelope proteins from a Guangdo
A;Reference number: PC4407
A;Accession: PC4407
A;Molecule type: genomic RNA
A;Residues: 1-415 <LIA>
A;Cross-references: UNIPROT:Q7LZY4; UNIPARC:UPI0000178545
A;Note: the authors translated the codon ATA for residues 93 and 249 as Met
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: envelope protein

Query Match 60.7%; Score 37; DB 2; Length 415;
Best Local Similarity 41.7%; Pred. No. 40;
Matches 5; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

```
Qy          3 CXXGPXTWXCXP 14
              |   | :| | |
Db        329 CGIVPASWVCGP 340
```

RESULT 10
S06022
regulatory protein O2 - maize
C;Species: Zea mays (maize)
C;Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 31-Dec-2004
C;Accession: S06022; S06009
R;Hartings, H.; Maddaloni, M.; Lazzaroni, N.; di Fonzo, N.; Motto, M.; Salamini, F.; Tho
EMBO J. 8, 2795-2801, 1989
A;Title: The O2 gene which regulates zein deposition in maize endosperm encodes a protei
A;Reference number: S06022; MUID:90059860; PMID:2479535
A;Accession: S06022
A;Molecule type: mRNA
A;Residues: 1-460 <HAR>
A;Cross-references: UNIPROT:P12959; UNIPARC:UPI000016E05D; GB:X16618; NID:g22383; PIDN:C
R;Maddaloni, M.; di Fonzo, N.; Hartings, H.; Lazzaroni, N.; Salamini, F.; Thompson, R.;
Nucleic Acids Res. 17, 7532, 1989
A;Title: The sequence of the zein regulatory gene opaque-2 (02) of Zea Mays.
A;Reference number: S06009; MUID:90016825; PMID:2798113
A;Accession: S06009
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-22,29-149,'D',151-460 <MAD>
A;Cross-references: UNIPARC:UPI00001794F4; EMBL:X15544
C;Genetics:
A;Gene: opaque 2
A;Map position: 7
A;Introns: 148/3; 168/3; 238/2; 263/3; 305/3
C;Superfamily: BZIP protein; fos/jun DNA-binding domain homology
C;Keywords: DNA binding; nucleus; transcription regulation
F;227-267/Domain: fos/jun DNA-binding domain homology <FJD>

Query Match 60.7%; Score 37; DB 2; Length 460;
Best Local Similarity 71.4%; Pred. No. 43;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          6 GPXTWXC 12
              || || |
Db        436 GPYTWTC 442
```

RESULT 11
H84137
hypothetical protein BH3904 [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C;Accession: H84137
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132

```
A;Accession: H84137
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-475 <STO>
A;Cross-references: UNIPROT:Q9K628; UNIPARC:UPI00000C432F; GB:AP001520; GB:BA000004; NID
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH3904

  Query Match             60.7%;  Score 37;  DB 2;  Length 475;
  Best Local Similarity   62.5%;  Pred. No. 45;
  Matches    5;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          3 CXXGPXTW 10
              |  || ||
Db        156 CAGGPSTW 163


RESULT 12
S25265
outer membrane protein irgA precursor - Vibrio cholerae
N;Alternate names: ferrienterochelin receptor homolog
C;Species: Vibrio cholerae
C;Date: 28-May-1993 #sequence_revision 28-May-1993 #text_change 09-Jul-2004
C;Accession: S25265; A37834
R;Goldberg, M.B.; Boyko, S.A.; Butterton, J.R.; Stoebner, J.A.; Payne, S.M.; Calderwood,
Mol. Microbiol. 6, 2407-2418, 1992
A;Title: Characterization of a Vibrio cholerae virulence factor homologous to the family
A;Reference number: S25265; MUID:93023868; PMID:1406279
A;Accession: S25265
A;Molecule type: DNA
A;Residues: 1-652 <GOL>
A;Cross-references: UNIPROT:P27772; UNIPARC:UPI0000148DB5; GB:U72152; EMBL:M63192; NID:g
A;Note: the sequence from Fig. 3 is inconsistent with that from Fig. 2 in having 299-Thr
R;Goldberg, M.B.; Boyko, S.A.; Calderwood, S.B.
J. Bacteriol. 172, 6863-6870, 1990
A;Title: Transcriptional regulation by iron of a Vibrio cholerae virulence gene and homo
A;Reference number: A37834; MUID:91072235; PMID:2174861
A;Accession: A37834
A;Molecule type: DNA
A;Residues: 1-152,'D' <GO2>
A;Cross-references: UNIPARC:UPI000017838A; GB:M37773
C;Genetics:
A;Gene: irgA
C;Superfamily: ferrienterochelin receptor; tonB-dependent receptor amino-terminal homolo
C;Keywords: membrane protein
F;1-25/Domain: signal sequence #status predicted <SIG>
F;26-652/Product: outer membrane protein irgA #status predicted <MAT>
F;68-214/Domain: tonB-dependent receptor amino-terminal homology <TNN>
F;367-652/Domain: tonB-dependent receptor carboxyl-terminal homology <TNC>

  Query Match             60.7%;  Score 37;  DB 2;  Length 652;
  Best Local Similarity   41.7%;  Pred. No. 59;
  Matches    5;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          3 CXXGPXTWXCXP 14
              |  ||  |   |
Db        492 CTAGPNQWGATP 503


RESULT 13
D82317
iron-regulated outer membrane virulence protein, TonB receptor family VC0475 [imported]
C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: D82317
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A;Reference number: A82035; MUID:20406833; PMID:10952301
A;Accession: D82317
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-652 <HEI>
A;Cross-references: UNIPROT:P27772; UNIPARC:UPI000012D88F; GB:AE004134; GB:AE003852; NID:
A;Experimental source: serogroup O1; strain N16961; biotype El Tor
C;Genetics:
A;Gene: VC0475
A;Map position: 1
C;Superfamily: ferrienterochelin receptor; tonB-dependent receptor amino-terminal homolog

  Query Match             60.7%;  Score 37;  DB 2;  Length 652;
  Best Local Similarity   41.7%;  Pred. No. 59;
  Matches    5;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          3 CXXGPXTWXCXP 14
              |  ||  |   |
Db        492 CTAGPNQWGAAP 503


RESULT 14
RRWVEV
genome polyprotein - equine arteritis virus
N;Contains: RNA-directed RNA polymerase (EC 2.7.7.48)
C;Species: equine arteritis virus
A;Note: host Equus caballus (domestic horse)
C;Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 09-Jul-2004
C;Accession: A39925; S10158; B39925
R;Den Boon, J.A.; Snijder, E.J.; Chirnside, E.D.; De Vries, A.A.F.; Horzinek, M.C.; Spaar
J. Virol. 65, 2910-2920, 1991
A;Title: Equine arteritis virus is not a togavirus but belongs to the coronaviruslike sup
A;Reference number: A39925; MUID:91237805; PMID:1851863
A;Accession: A39925
A;Molecule type: genomic RNA
A;Residues: 1-3175 <DEN>
A;Cross-references: UNIPROT:P19811; UNIPARC:UPI0000134685; EMBL:X53459
A;Note: a -1 ribosomal frameshift occurs between the codons AAC for 1727-Asn and CUG for
R;de Vries, A.A.F.; Chirnside, E.D.; Bredenbeek, P.J.; Gravestein, L.A.; Horzinek, M.C.;
Nucleic Acids Res. 18, 3241-3247, 1990
A;Title: All subgenomic mRNAs of equine arteritis virus contain a common leader sequence
A;Reference number: S10158; MUID:90287699; PMID:2162519
A;Accession: S10158
A;Status: translation not shown
A;Molecule type: genomic RNA
A;Residues: 1-17 <VRI>
A;Cross-references: UNIPARC:UPI0000172725; EMBL:X52277
C;Superfamily: equine arteritis virus RNA-directed RNA polymerase
C;Keywords: nucleotidyltransferase

  Query Match             60.7%;  Score 37;  DB 1;  Length 3175;
  Best Local Similarity   35.7%;  Pred. No. 2.3e+02;
  Matches    5;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          1 YXCXXGPXTWXCXP 14
              | |       | | |
Db        242 YVCDISEADWSCLP 255


RESULT 15
I52427
guanine-nucleotide-releasing protein Mss4 - human
C;Species: Homo sapiens (man)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C;Accession: I52427
R;Yu, H.; Schreiber, S.L.
Biochemistry 34, 9103-9110, 1995
A;Title: Cloning, Zn2+ binding, and structural characterization of the guanine nucleotide
A;Reference number: I52427; MUID:95345082; PMID:7619808
A;Accession: I52427
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-123 <RES>
A;Cross-references: UNIPROT:P47224; UNIPARC:UPI00001117CC; GB:S78873; NID:g1037135; PIDN
C;Genetics:
```

A;Gene: GDB:MSS4
A;Cross-references: GDB:683578

Query Match 59.0%; Score 36; DB 2; Length 123;
Best Local Similarity 50.0%; Pred. No. 21;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy          3 CXXGPXTWXC 12
              |  ||  | |
Db         97 CEIGPIGWHC 106
```

Search completed: March 31, 2006, 16:37:15
Job time : 8.70647 secs

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: March 31, 2006, 16:09:36 ; Search time 52.4478 Seconds
(without alignments)
188.328 Million cell updates/sec

Title: US-10-609-217-85
Perfect score: 61
Sequence: 1 YXCXXGPXTWXCXP 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	46	75.4	352	2	Q4IMN3_GIBZE	Q4imn3 gibberella
2	45	73.8	19	1	DURC_STRGP	P36503 streptomyce
3	45	73.8	19	1	LANC_STRS6	P38655 streptomyce
4	43	70.5	378	1	TAZ_DROME	Q9v6g5 drosophila
5	43	70.5	414	2	Q4SAV9_TETNG	Q4sav9 tetraodon n
6	41	67.2	532	2	Q8WWV6_HUMAN	Q8wwv6 homo sapien
7	41	67.2	534	2	Q96SA2_HUMAN	Q96sa2 homo sapien
8	41	67.2	577	2	Q5REH9_PONPY	Q5reh9 pongo pygma
9	41	67.2	589	2	Q5R770_PONPY	Q5r770 pongo pygma
10	40.5	66.4	499	2	Q6ARY7_DESPS	Q6ary7 desulfotale
11	40	65.6	167	2	Q6ZW93_HUMAN	Q6zw93 homo sapien
12	40	65.6	173	2	Q5VHX3_EAV	Q5vhx3 equine arte
13	40	65.6	180	2	Q4I355_GIBZE	Q4i355 gibberella
14	40	65.6	345	2	Q7JKP2_CAEEL	Q7jkp2 caenorhabdi
15	40	65.6	358	2	Q9U362_CAEEL	Q9u362 caenorhabdi
16	40	65.6	698	2	Q810G8_RAT	Q810g8 rattus norv
17	40	65.6	1523	2	Q810G9_RAT	Q810g9 rattus norv
18	40	65.6	1525	2	Q5F364_CHICK	Q5f364 gallus gall
19	40	65.6	1528	1	MRP1_MOUSE	O35379 mus musculu
20	40	65.6	1530	2	Q8HXQ5_BOVIN	Q8hxq5 bos taurus
21	40	65.6	1531	2	Q6UR05_CANFA	Q6ur05 canis famil
22	40	65.6	1531	2	Q864R9_MACFA	Q864r9 macaca fasc
23	40	65.6	1531	2	Q864S0_MACFA	Q864s0 macaca fasc
24	40	65.6	1532	2	Q810E4_RAT	Q810e4 rattus norv
25	40	65.6	1532	2	Q8CG09_RAT	Q8cg09 rattus norv
26	40	65.6	2022	2	Q61J27_CAEBR	Q61j27 caenorhabdi
27	39	63.9	172	2	Q6ZT75_HUMAN	Q6zt75 homo sapien
28	39	63.9	172	2	Q6ZWC2_HUMAN	Q6zwc2 homo sapien
29	39	63.9	173	2	Q9WD22_EAV	Q9wd22 equine arte
30	39	63.9	373	2	O70709_9POXV	O70709 anomala cup
31	39	63.9	691	2	O43333_HUMAN	O43333 homo sapien
32	39	63.9	1215	2	Q68CP7_HUMAN	Q68cp7 homo sapien
33	39	63.9	1400	2	Q9UQ98_HUMAN	Q9uq98 homo sapien
34	39	63.9	1439	2	Q59GI9_HUMAN	Q59gi9 homo sapien
35	39	63.9	1456	2	Q9UQA0_HUMAN	Q9uqa0 homo sapien
36	39	63.9	1459	2	Q9UQ97_HUMAN	Q9uq97 homo sapien
37	39	63.9	1515	2	Q9UQ99_HUMAN	Q9uq99 homo sapien
38	39	63.9	1531	1	MRP1_HUMAN	P33527 homo sapien
39	38	62.3	61	2	O70227_RAT	O70227 rattus norv
40	38	62.3	167	2	Q651J0_ORYSA	Q651j0 oryza sativ
41	38	62.3	285	2	Q8NAV2_HUMAN	Q8nav2 homo sapien
42	38	62.3	329	2	Q7Z758_HUMAN	Q7z758 homo sapien
43	38	62.3	336	2	Q5PR99_BRARE	Q5pr99 brachydanio
44	38	62.3	967	2	Q59FS0_HUMAN	Q59fs0 homo sapien
45	37.5	61.5	2465	2	Q4RXZ7_TETNG	Q4rxz7 tetraodon n

ALIGNMENTS

RESULT 1
Q4IMN3_GIBZE
ID Q4IMN3_GIBZE PRELIMINARY; PRT; 352 AA.
AC Q4IMN3;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=FG01525.1;
OS Gibberella zeae PH-1.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
OX NCBI_TaxID=229533;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PH-1;
RA Birren B., Nusbaum C., Abouelleil A., Allen N., Anderson S.,
RA Arachchi H.M., Barna N., Bastien V., Bloom T., Boguslavkiy L.,
RA Boukhgalter B., Butler J., Calvo S.E., Camarata J., Chang J.,
RA Choepel Y., Collymore A., Cook A., Cooke P., Corum B., DeArellano K.,
RA Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
RA Erickson J., Faro S., Ferreira P., FitzGerald M., Gage D., Galagan J.,
RA Gardyna S., Gnerre S., Graham L., Grand-Pierre N., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Horton L., Hulme W., Iliev I.,
RA Jaffe D., Johnson R., Jones C., Kamal M., Kamat A., Karatas A.,
RA Kells C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
RA Ma L.-J., Mabbitt R., MacLean C., Macdonald P., Major J., Manning J.,
RA Matthews C., Mauceli E., McCarthy M., Meldrim J., Meneus L.,
RA Mihova T., Mlenga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
RA Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neil D.,
RA Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,
RA Rachupka A., Ramasamy U., Raymond C., Retta R., Rise C., Rogov P.,
RA Roman J., Schauer S., Schupback R., Seaman S., Severy P., Smirnov S.,
RA Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,
RA Talamas J., Tesfaye S., Theodore J., Topham K., Travers M.,
RA Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,
RA Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
RA Lander E.;
RT "Fusarium graminearum genome sequence.";
RL Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AACM01000077; EAA68151.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 352 AA; 38308 MW; 670BA49FC645A7F8 CRC64;

Query Match 75.4%; Score 46; DB 2; Length 352;
Best Local Similarity 50.0%; Pred. No. 4.4;
Matches 6; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
Qy          3 CXXGPXTWXCXP 14
              |  | || | |
Db        184 CTSNPSTWRCYP 195
```

```
RESULT 2
DURC_STRGP
ID   DURC_STRGP     STANDARD;      PRT;    19 AA.
AC   P36503;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Lantibiotic duramycin C.
OS   Streptomyces griseoluteus.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=29306;
RN   [1]
RP   PROTEIN SEQUENCE.
RC   STRAIN=R2107;
RX   MEDLINE=91107436; PubMed=2125590;
RA   Fredenhagen A., Fendrich G., Marki F., Marki W., Gruner J.,
RA   Raschdorf F., Peter H.H.;
RT   "Duramycins B and C, two new lanthionine containing antibiotics as
RT   inhibitors of phospholipase A2. Structural revision of duramycin and
RT   cinnamycin.";
RL   J. Antibiot. 43:1403-1412(1990).
RN   [2]
RP   STRUCTURE BY NMR.
RA   Zimmermann N., Freund S., Fredenhagen A., Jung G.;
RT   "Solution structure of the lantibiotics duramycin B and C.";
RL   (In) Schneider C.H., Eberles A.N. (eds.);
RL   Peptides 1992, pp.519-520, Escom Science Publishers, Leiden (1993).
RN   [3]
RP   STRUCTURE BY NMR.
RX   MEDLINE=93387292; PubMed=8375380;
RA   Zimmermann N., Freund S., Fredenhagen A., Jung G.;
RT   "Solution structures of the lantibiotics duramycin B and C.";
RL   Eur. J. Biochem. 216:419-428(1993).
CC   -!- FUNCTION: Acts as inhibitor of phospholipase A2.
CC   -!- PTM: Maturation of lantibiotics involves the enzymic conversion of
CC       Thr, and Ser into dehydrated AA and the formation of thioether
CC       bonds with cysteine or the formation of dialkylamine bonds with
CC       lysine. This is followed by membrane translocation and cleavage of
CC       the modified precursor.
CC   -!- SIMILARITY: Belongs to the type B lantibiotic family.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
KW   Antibiotic; Antimicrobial; Bacteriocin; Direct protein sequencing;
KW   Lantibiotic; Thioether bond.
FT   CROSSLNK      1     18       Beta-methyllanthionine (Cys-Thr).
FT   CROSSLNK      4     14       Lanthionine (Ser-Cys).
FT   CROSSLNK      5     11       Beta-methyllanthionine (Cys-Thr).
FT   CROSSLNK      6     19       Lysinoalanine (Ser-Lys).
SQ   SEQUENCE   19 AA;  2007 MW;  E2404ECE3F95286A CRC64;

  Query Match             73.8%;  Score 45;  DB 1;  Length 19;
  Best Local Similarity   60.0%;  Pred. No. 0.44;
  Matches    6;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          3 CXXGPXTWXC 12
              |  || || |
Db          5 CSYGPLTWSC 14


RESULT 3
LANC_STRS6
ID   LANC_STRS6     STANDARD;      PRT;    19 AA.
AC   P38655;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Lantibiotic ancovenin.
OS   Streptomyces sp. (strain A647P-2).
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=72591;
RN   [1]
RP   PROTEIN SEQUENCE.
RA   Wakamiya T., Ueki Y., Shiba T., Kido Y., Motoki Y.;
RT   "The structure of ancovenin, a new peptide inhibitor of angiotensin I
RT   converting enzyme.";
RL   Tetrahedron Lett. 26:665-668(1985).
CC   -!- FUNCTION: Acts as an inhibitor of angiotensin I converting enzyme.
CC   -!- PTM: Maturation of lantibiotics involves the enzymic conversion of
CC       Thr, and Ser into dehydrated AA and the formation of thioether
CC       bonds with cysteine or the formation of dialkylamine bonds with
CC       lysine. This is followed by membrane translocation and cleavage of
CC       the modified precursor.
CC   -!- SIMILARITY: Belongs to the type B lantibiotic family.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   PIR; A61284; EWSMAN.
KW   Antibiotic; Antimicrobial; Bacteriocin; Direct protein sequencing;
KW   Lantibiotic; Thioether bond.
FT   CROSSLNK      1     18       Beta-methyllanthionine (Cys-Thr).
FT   CROSSLNK      4     14       Lanthionine (Ser-Cys).
FT   CROSSLNK      5     11       Beta-methyllanthionine (Cys-Thr).
FT   CROSSLNK      6     19       Lysinoalanine (Ser-Lys).
SQ   SEQUENCE   19 AA;  2033 MW;  F434299E2736286A CRC64;

  Query Match             73.8%;  Score 45;  DB 1;  Length 19;
  Best Local Similarity   60.0%;  Pred. No. 0.44;
  Matches    6;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          3 CXXGPXTWXC 12
              |  || || |
Db          5 CSFGPLTWSC 14


RESULT 4
TAZ_DROME
ID   TAZ_DROME      STANDARD;      PRT;   378 AA.
AC   Q9V6G5; Q8ML32; Q8SZ79; Q9U9U8; Q9V6G4;
DT   28-FEB-2003 (Rel. 41, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Tafazzin homolog.
GN   Name=tafazzin; ORFNames=CG8766;
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   NUCLEOTIDE SEQUENCE (ISOFORM A).
RC   STRAIN=Canton-S;
RA   Benevolenskaya E.V., Frolov M.V., Birchler J.A.;
RT   "Drosophila homolog of the human G4.5 gene encoding tafazzin
RT   proteins.";
RL   Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC   STRAIN=Berkeley;
RX   MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
```

```
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [3]
RP   GENOME REANNOTATION, AND ALTERNATIVE SPLICING.
RX   MEDLINE=22426069; PubMed=12537572;
RA   Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA   Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA   Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA   Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA   Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA   Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA   Lewis S.E.;
RT   "Annotation of the Drosophila melanogaster euchromatic genome: a
RT   systematic review.";
RL   Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN   [4]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM A).
RC   STRAIN=Berkeley; TISSUE=Embryo;
RX   MEDLINE=22426066; PubMed=12537569;
RA   Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,
RA   George R.A., Guarin H., Kronmiller B., Pacleb J.M., Park S., Wan K.H.,
RA   Rubin G.M., Celniker S.E.;
RT   "A Drosophila full-length cDNA resource.";
RL   Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).
RN   [5]
RP   NUCLEOTIDE SEQUENCE OF 1-39 (ISOFORM A).
RC   STRAIN=Canton-S;
RX   MEDLINE=20556141; PubMed=11102369;
RA   Frolov M.V., Benevolenskaya E.V., Birchler J.A.;
RT   "The oxen gene of Drosophila encodes a homolog of subunit 9 of yeast
RT   ubiquinol-cytochrome c oxidoreductase complex: evidence for modulation
RT   of gene expression in response to mitochondrial activity.";
RL   Genetics 156:1727-1736(2000).
CC   -!- INTERACTION:
CC       Q9VA96:CG15530; NbExp=1; IntAct=EBI-244104, EBI-151361;
CC   -!- SUBCELLULAR LOCATION: Isoforms with hydrophobic N-terminus are
CC       thought to be membrane-anchored (By similarity).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=3;
CC       Name=A;
CC         IsoId=Q9V6G5-1; Sequence=Displayed;
CC       Name=B;
CC         IsoId=Q9V6G5-2; Sequence=VSP_004450;
CC         Note=No experimental confirmation available;
CC       Name=C;
CC         IsoId=Q9V6G5-3; Sequence=VSP_007017;
CC         Note=No experimental confirmation available;
CC   -!- DOMAIN: The hydrophilic domain may serve as an exposed loop
CC       interacting with other proteins (By similarity).
CC   -!- SIMILARITY: Belongs to the taffazin family.
CC   -!- CAUTION: Ref.1 sequence differs from that shown due to a
CC       frameshift in position 117.
CC   ---------------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------------
DR   EMBL; AF148684; AAD48409.1; ALT_FRAME; mRNA.
DR   EMBL; AE003821; AAF58461.2; -; Genomic_DNA.
DR   EMBL; AE003821; AAF58462.3; -; Genomic_DNA.
DR   EMBL; AE003821; AAM68652.2; -; Genomic_DNA.
DR   EMBL; AY071059; AAL48681.1; -; mRNA.
DR   EMBL; AF017783; -; NOT_ANNOTATED_CDS; Genomic_DNA.
DR   IntAct; Q9V6G5; -.
DR   Ensembl; CG8766; Drosophila melanogaster.
DR   FlyBase; FBgn0026619; tafazzin.
DR   GO; GO:0016021; C:integral to membrane; NAS.
DR   InterPro; IPR002123; Acyltransferase.
DR   InterPro; IPR000872; Tafazzin.
DR   PANTHER; PTHR12497; Tafazzin; 1.
DR   Pfam; PF01553; Acyltransferase; 1.
DR   PRINTS; PR00979; TAFAZZIN.
KW   Alternative splicing; Transmembrane.
FT   TRANSMEM    138    158       Potential.
FT   REGION      243    283       Hydrophilic (By similarity).
FT   VARSPLIC      1    100       Missing (in isoform B).
FT                               /FTId=VSP_004450.
FT   VARSPLIC      1     41       MFMVVCSNLRRPGHVGAASAARNINWLISEGYTPPIRAMAR
FT                               -> M (in isoform C).
FT                               /FTId=VSP_007017.
SQ   SEQUENCE   378 AA;  43016 MW;  88690B6E2731FFB9 CRC64;

  Query Match             70.5%;  Score 43;  DB 1;  Length 378;
  Best Local Similarity   42.9%;  Pred. No. 16;
  Matches    6;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          1 YXCXXGPXTWXCXP 14
              | |   |  | | |
Db        189 YSCFDDPGLWGCLP 202


RESULT 5
Q4SAV9_TETNG
ID   Q4SAV9_TETNG PRELIMINARY;      PRT;   414 AA.
AC   Q4SAV9;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Chromosome 3 SCAF14679, whole genome shotgun sequence.
GN   ORFNames=GSTENG00021242001;
OS   Tetraodon nigroviridis (Green puffer).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC   Tetradontoidea; Tetraodontidae; Tetraodon.
OX   NCBI_TaxID=99883;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA   Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA   Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA   Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
```

```
RA   Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA   Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA   Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA   Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA   Kellis M., Volff JN., Guigo R., Zody M.C., Mesirov J.,
RA   Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA   Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA   Wincker P., Lander E.S., Weissenbach J., Roest Crollius H.;
RT   "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT   the early vertebrate proto-karyotype.";
RL   Nature 431:946-957(2004).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RG   Genoscope; Whitehead Institute Centre for Genome Research;
RL   Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; CAAE01014679; CAG02223.1; -; Genomic_DNA.
SQ   SEQUENCE   414 AA;  45368 MW;  0522D03EA381377E CRC64;

  Query Match             70.5%;  Score 43;  DB 2;  Length 414;
  Best Local Similarity   50.0%;  Pred. No. 17;
  Matches    6;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          3 CXXGPXTWXCXP 14
              |   | || | |
Db        155 CRMSPSTWGCCP 166


RESULT 6
Q8WWV6_HUMAN
ID   Q8WWV6_HUMAN PRELIMINARY;      PRT;   532 AA.
AC   Q8WWV6;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Fc alpha/mu receptor.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=21638011; PubMed=11779189;
RA   McDonald K.J., Cameron A.J.M., Allen J.M., Jardine A.G.;
RT   "Expression of Fc alpha/mu receptor by human mesangial cells: a
RT   candidate receptor for immune complex deposition in IgA nephropathy.";
RL   Biochem. Biophys. Res. Commun. 290:438-442(2002).
DR   EMBL; AY063125; AAL51154.1; -; mRNA.
DR   Ensembl; ENSG00000162897; Homo sapiens.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR007110; Ig-like.
DR   SMART; SM00409; IG; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
KW   Immunoglobulin domain; Receptor.
SQ   SEQUENCE   532 AA;  57144 MW;  D347A23C0F41EED3 CRC64;

  Query Match             67.2%;  Score 41;  DB 2;  Length 532;
  Best Local Similarity   50.0%;  Pred. No. 50;
  Matches    6;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 YXCXXGPXTWXC 12
              | |  ||  | |
Db         96 YWCRLGPPRWIC 107


RESULT 7
Q96SA2_HUMAN
ID   Q96SA2_HUMAN PRELIMINARY;      PRT;   534 AA.
AC   Q96SA2;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   FKSG87 protein.
GN   Name=FKSG87;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Wang Y.-G., Gong L.;
RL   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF354295; AAK39522.1; -; mRNA.
DR   Ensembl; ENSG00000162897; Homo sapiens.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR007110; Ig-like.
DR   SMART; SM00409; IG; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
KW   Immunoglobulin domain.
SQ   SEQUENCE   534 AA;  56749 MW;  6EF8050E412AF91C CRC64;

  Query Match             67.2%;  Score 41;  DB 2;  Length 534;
  Best Local Similarity   50.0%;  Pred. No. 50;
  Matches    6;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 YXCXXGPXTWXC 12
              | |  ||  | |
Db        116 YWCRLGPPRWIC 127


RESULT 8
Q5REH9_PONPY
ID   Q5REH9_PONPY PRELIMINARY;      PRT;   577 AA.
AC   Q5REH9;
DT   01-FEB-2005 (TrEMBLrel. 29, Created)
DT   01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT   01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE   Hypothetical protein DKFZp469K1129.
GN   Name=DKFZp469K1129;
OS   Pongo pygmaeus (Orangutan).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Pongo.
OX   NCBI_TaxID=9600;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Kidney;
RG   The German cDNA Consortium;
RA   Ottenwaelder B., Obermaier B., Deutschenbaur S., Schaipp A.,
RA   Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL   Submitted (NOV-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; CR857549; CAH89828.1; -; mRNA.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR007110; Ig-like.
DR   SMART; SM00409; IG; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
KW   Hypothetical protein; Immunoglobulin domain.
SQ   SEQUENCE   577 AA;  62062 MW;  AA0FCEE7AB9C4BCD CRC64;

  Query Match             67.2%;  Score 41;  DB 2;  Length 577;
  Best Local Similarity   50.0%;  Pred. No. 54;
  Matches    6;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 YXCXXGPXTWXC 12
              | |  ||  | |
Db        129 YWCRLGPPRWIC 140


RESULT 9
Q5R770_PONPY
ID   Q5R770_PONPY PRELIMINARY;      PRT;   589 AA.
```

```
AC   Q5R770;
DT   01-FEB-2005 (TrEMBLrel. 29, Created)
DT   01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT   01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE   Hypothetical protein DKFZp469A0319.
GN   Name=DKFZp469A0319;
OS   Pongo pygmaeus (Orangutan).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Pongo.
OX   NCBI_TaxID=9600;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Kidney;
RG   The German cDNA Consortium;
RA   Poustka A., Albert R., Moosmayer P., Schupp I., Wellenreuther R.,
RA   Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL   Submitted (NOV-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; CR860248; CAH92390.1; -; mRNA.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR007110; Ig-like.
DR   SMART; SM00409; IG; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
KW   Hypothetical protein; Immunoglobulin domain.
SQ   SEQUENCE   589 AA;  63435 MW;  255BF0FEAACCA812 CRC64;

  Query Match             67.2%;  Score 41;  DB 2;  Length 589;
  Best Local Similarity   50.0%;  Pred. No. 55;
  Matches    6;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 YXCXXGPXTWXC 12
              | |  ||  | |
Db        141 YWCRLGPPRWIC 152


RESULT 10
Q6ARY7_DESPS
ID   Q6ARY7_DESPS PRELIMINARY;      PRT;   499 AA.
AC   Q6ARY7;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Related to cytochrome-c3 hydrogenase (NiFeSe), large subunit.
GN   OrderedLocusNames=DP0159;
OS   Desulfotalea psychrophila.
OC   Bacteria; Proteobacteria; Deltaproteobacteria; Desulfobacterales;
OC   Desulfobulbaceae; Desulfotalea.
OX   NCBI_TaxID=84980;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=LSv54 / DSM 12343;
RX   PubMed=15305914; DOI=10.1111/j.1462-2920.2004.00665.x;
RA   Rabus R., Ruepp A., Frickey T., Rattei T., Fartmann B., Stark M.,
RA   Bauer M., Zibat A., Lombardot T., Becker I., Amann J., Gellner K.,
RA   Teeling H., Leuschner W.D., Gloeckner F.-O., Lupas A.N., Amann R.,
RA   Klenk H.-P.;
RT   "The genome of Desulfotalea psychrophila, a sulfate-reducing bacterium
RT   from permanently cold Arctic sediments.";
RL   Environ. Microbiol. 6:887-902(2004).
DR   EMBL; CR522870; CAG34888.1; -; Genomic_DNA.
DR   GO; GO:0008901; F:ferredoxin hydrogenase activity; IEA.
DR   GO; GO:0046872; F:metal ion binding; IEA.
DR   GO; GO:0016151; F:nickel ion binding; IEA.
DR   GO; GO:0016491; F:oxidoreductase activity; IEA.
DR   GO; GO:0006118; P:electron transport; IEA.
DR   InterPro; IPR001501; Ni_hdL.
DR   Pfam; PF00374; NiFeSe_Hases; 1.
DR   PROSITE; PS00507; NI_HGENASE_L_1; 1.
DR   PROSITE; PS00508; NI_HGENASE_L_2; 1.
KW   Complete proteome; Metal-binding; Nickel; Oxidoreductase.
SQ   SEQUENCE   499 AA;  55328 MW;  8DC870ABF65B7618 CRC64;

  Query Match             66.4%;  Score 40.5;  DB 2;  Length 499;
  Best Local Similarity   50.0%;  Pred. No. 58;
  Matches    7;  Conservative    0;  Mismatches    6;  Indels    1;  Gaps    1;

Qy          1 YXCXXGPXTWXCXP 14
              | |   | || | |
Db        443 YECIV-PTTWNCSP 455


RESULT 11
Q6ZW93_HUMAN
ID   Q6ZW93_HUMAN PRELIMINARY;      PRT;   167 AA.
AC   Q6ZW93;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Hypothetical protein FLJ41423.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Hippocampus;
RA   Kawakami B., Sugiyama A., Takemoto M., Sugiyama T., Irie R.,
RA   Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA   Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA   Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,
RA   Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A., Suzuki Y.,
RA   Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
RL   Submitted (JUL-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK123417; BAC85611.1; -; mRNA.
SQ   SEQUENCE   167 AA;  17960 MW;  266132D59393C276 CRC64;

  Query Match             65.6%;  Score 40;  DB 2;  Length 167;
  Best Local Similarity   50.0%;  Pred. No. 26;
  Matches    5;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          3 CXXGPXTWXC 12
              |  ||  | |
Db         83 CRQGPSVWSC 92


RESULT 12
Q5VHX3_EAV
ID   Q5VHX3_EAV  PRELIMINARY;      PRT;   173 AA.
AC   Q5VHX3;
DT   01-FEB-2005 (TrEMBLrel. 29, Created)
DT   01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT   01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE   Large envelope protein (Fragment).
GN   Name=ORF5;
OS   Equine arteritis virus (EAV).
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC   Arteriviridae; Arterivirus.
OX   NCBI_TaxID=11047;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=S4;
RA   Mittelholzer C., Johansson I., Baule C., Hannant D., Paton D.,
RA   Autorino G.L., Nowotny N., Belak S.;
RT   "Extended phylogeny of equine arteritis virus: division into new
RT   subgroups.";
RL   Submitted (OCT-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY453342; AAS17004.1; -; Genomic_RNA.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   InterPro; IPR001332; Arteri_glycop.
DR   InterPro; IPR003241; EAV_ORF5.
DR   Pfam; PF00951; Arteri_Gl; 1.
DR   ProDom; PD002371; EAV_ORF5; 1.
KW   Envelope protein.
FT   NON_TER       1      1
FT   NON_TER     173    173
```

```
SQ   SEQUENCE   173 AA;  19488 MW;  9147CBD1DD750ADE CRC64;

  Query Match             65.6%;  Score 40;  DB 2;  Length 173;
  Best Local Similarity   41.7%;  Pred. No. 27;
  Matches    5;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 YXCXXGPXTWXC 12
              | |   |  | ||
Db          5 YNCSASPTCWYC 16


RESULT 13
Q4I355_GIBZE
ID   Q4I355_GIBZE PRELIMINARY;      PRT;   180 AA.
AC   Q4I355;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Predicted protein.
GN   ORFNames=FG08353.1;
OS   Gibberella zeae PH-1.
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC   Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
OX   NCBI_TaxID=229533;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=PH-1;
RA   Birren B., Nusbaum C., Abouelleil A., Allen N., Anderson S.,
RA   Arachchi H.M., Barna N., Bastien V., Bloom T., Boguslavkiy L.,
RA   Boukhgalter B., Butler J., Calvo S.E., Camarata J., Chang J.,
RA   Choepel Y., Collymore A., Cook A., Cooke P., Corum B., DeArellano K.,
RA   Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
RA   Erickson J., Faro S., Ferreira P., FitzGerald M., Gage D., Galagan J.,
RA   Gardyna S., Gnerre S., Graham L., Grand-Pierre N., Hafez N.,
RA   Hagopian D., Hagos B., Hall J., Horton L., Hulme W., Iliev I.,
RA   Jaffe D., Johnson R., Jones C., Kamal M., Kamat A., Karatas A.,
RA   Kells C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
RA   Ma L.-J., Mabbitt R., MacLean C., Macdonald P., Major J., Manning J.,
RA   Matthews C., Mauceli E., McCarthy M., Meldrim J., Meneus L.,
RA   Mihova T., Mlenga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
RA   Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neil D.,
RA   Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,
RA   Rachupka A., Ramasamy U., Raymond C., Retta R., Rise C., Rogov P.,
RA   Roman J., Schauer S., Schupback R., Seaman S., Severy P., Smirnov S.,
RA   Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,
RA   Talamas J., Tesfaye S., Theodore J., Topham K., Travers M.,
RA   Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,
RA   Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
RA   Lander E.;
RT   "Fusarium graminearum genome sequence.";
RL   Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; AACM01000335; EAA72141.1; -; Genomic_DNA.
SQ   SEQUENCE   180 AA;  20463 MW;  94C7B5242FEE6ED9 CRC64;

  Query Match             65.6%;  Score 40;  DB 2;  Length 180;
  Best Local Similarity   41.7%;  Pred. No. 28;
  Matches    5;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 YXCXXGPXTWXC 12
              : |   ||  | |
Db         99 HNCSFGPAPWEC 110


RESULT 14
Q7JKP2_CAEEL
ID   Q7JKP2_CAEEL PRELIMINARY;      PRT;   345 AA.
AC   Q7JKP2;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Hypothetical protein T22H2.6b.
GN   ORFNames=T22H2.6, T22H2.6B;
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC   STRAIN=Bristol N2;
RX   MEDLINE=99069613; PubMed=9851916;
RG   The C. elegans sequencing consortium;
RT   "Genome sequence of the nematode C. elegans: a platform for
RT   investigating biology.";
RL   Science 282:2012-2018(1998).
DR   EMBL; Z81595; CAB54305.1; -; Genomic_DNA.
DR   Ensembl; T22H2.6; Caenorhabditis elegans.
DR   WormBase; WBGene00011936; T22H2.6.
DR   WormPep; T22H2.6b; CE24005.
DR   InterPro; IPR000118; Granulin.
DR   Pfam; PF00396; Granulin; 2.
DR   SMART; SM00277; GRAN; 3.
DR   PROSITE; PS00799; GRANULINS; 2.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   345 AA;  38122 MW;  D93C75167C3650B9 CRC64;

  Query Match             65.6%;  Score 40;  DB 2;  Length 345;
  Best Local Similarity   50.0%;  Pred. No. 50;
  Matches    6;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          3 CXXGPXTWXCXP 14
              |  |  || | |
Db        111 CKLGDNTWGCCP 122


RESULT 15
Q9U362_CAEEL
ID   Q9U362_CAEEL PRELIMINARY;      PRT;   358 AA.
AC   Q9U362; Q9U361;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein T22H2.6a.
GN   ORFNames=T22H2.6, T22H2.6a;
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC   STRAIN=Bristol N2;
RX   MEDLINE=99069613; PubMed=9851916;
RG   The C. elegans sequencing consortium;
RT   "Genome sequence of the nematode C. elegans: a platform for
RT   investigating biology.";
RL   Science 282:2012-2018(1998).
DR   EMBL; Z81595; CAB54304.1; -; Genomic_DNA.
DR   PIR; T25137; T25137.
DR   PIR; T25138; T25138.
DR   HSSP; P28799; 1G26.
DR   Ensembl; T22H2.6; Caenorhabditis elegans.
DR   WormBase; WBGene00011936; T22H2.6.
DR   WormPep; T22H2.6a; CE24004.
DR   InterPro; IPR000118; Granulin.
DR   Pfam; PF00396; Granulin; 3.
DR   SMART; SM00277; GRAN; 3.
DR   PROSITE; PS00799; GRANULINS; 2.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   358 AA;  39754 MW;  2AD5B8F9B70D1595 CRC64;

  Query Match             65.6%;  Score 40;  DB 2;  Length 358;
  Best Local Similarity   50.0%;  Pred. No. 52;
  Matches    6;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;
```

```
Qy          3 CXXGPXTWXCXP 14
              |  |  || | |
Db        111 CKLGDNTWGCCP 122
```

Search completed: March 31, 2006, 16:35:05
Job time : 52.4478 secs

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: March 31, 2006, 16:09:06 ; Search time 38.5572 Seconds
(without alignments)
113.955 Million cell updates/sec

Title: US-10-609-217-124
Perfect score: 29
Sequence: 1 XXXGPXTWXX 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	29	100.0	14	4	AAM98806	Aam98806 Human pep
2	29	100.0	14	4	AAM98462	Aam98462 Human pep
3	29	100.0	17	9	ADU91962	Adu91962 EPO-R ago
4	29	100.0	17	9	ADY54197	Ady54197 Amino aci
5	29	100.0	20	2	AAY13704	Aay13704 Erythropo
6	29	100.0	20	2	AAY13696	Aay13696 Erythropo
7	29	100.0	20	2	AAY13650	Aay13650 Erythropo
8	29	100.0	20	2	AAY13728	Aay13728 Erythropo
9	29	100.0	20	2	AAY13688	Aay13688 Erythropo
10	29	100.0	20	2	AAY13687	Aay13687 Erythropo
11	29	100.0	20	2	AAY13705	Aay13705 Erythropo
12	29	100.0	20	2	AAY26368	Aay26368 Erythropo
13	29	100.0	20	2	AAY13672	Aay13672 Erythropo
14	29	100.0	20	2	AAY13706	Aay13706 Erythropo
15	29	100.0	20	2	AAY13679	Aay13679 Erythropo
16	29	100.0	20	2	AAW27019	Aaw27019 Monomer s
17	29	100.0	20	2	AAW27001	Aaw27001 Monomer s
18	29	100.0	20	2	AAW26993	Aaw26993 Monomer s
19	29	100.0	20	2	AAW27020	Aaw27020 Monomer s
20	29	100.0	20	2	AAW27010	Aaw27010 Monomer s
21	29	100.0	20	2	AAW27018	Aaw27018 Monomer s
22	29	100.0	20	2	AAW27002	Aaw27002 Monomer s
23	29	100.0	20	2	AAW26986	Aaw26986 Monomer s
24	29	100.0	20	5	AAU74480	Aau74480 Human ery
25	29	100.0	20	7	ADD12273	Add12273 PDZ ligan
26	29	100.0	20	8	ADO20705	Ado20705 PSD-95 PD
27	29	100.0	20	8	ADO20840	Ado20840 NMDA rece
28	29	100.0	20	8	ADO20848	Ado20848 NMDA rece
29	29	100.0	20	8	ADO20867	Ado20867 NMDA rece
30	29	100.0	20	8	ADO20872	Ado20872 NMDA rece
31	29	100.0	20	8	ADO20769	Ado20769 PSD-95 PD
32	29	100.0	20	8	ADO20819	Ado20819 NMDA rece
33	29	100.0	20	8	ADO20824	Ado20824 NMDA rece
34	29	100.0	20	8	ADO20834	Ado20834 NMDA rece
35	29	100.0	20	8	ADO20844	Ado20844 NMDA rece
36	29	100.0	20	8	ADO20856	Ado20856 NMDA rece
37	29	100.0	20	8	ADO20860	Ado20860 NMDA rece
38	29	100.0	20	8	ADO20682	Ado20682 PSD-95 PD
39	29	100.0	20	8	ADO20821	Ado20821 NMDA rece
40	29	100.0	20	8	ADO20831	Ado20831 NMDA rece
41	29	100.0	20	8	ADO20833	Ado20833 NMDA rece
42	29	100.0	20	8	ADO20837	Ado20837 NMDA rece
43	29	100.0	20	8	ADO20851	Ado20851 NMDA rece
44	29	100.0	20	8	ADO20816	Ado20816 NMDA rece
45	29	100.0	20	8	ADO20823	Ado20823 NMDA rece

ALIGNMENTS

RESULT 1
AAM98806
ID AAM98806 standard; peptide; 14 AA.
XX
AC AAM98806;
XX
DT 24-JAN-2002 (first entry)
XX
DE Human peptide #2081 encoded by a SNP oligonucleotide.
XX
KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW complement related protein; cytochrome; kinesin; cytokine; interferon;
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;
KW nervous system disease.
XX
OS Homo sapiens.
XX
PN WO200147944-A2.
XX
PD 05-JUL-2001.
XX
PF 28-DEC-2000; 2000WO-US035498.
XX
PR 28-DEC-1999; 99US-0173419P.
PR 27-DEC-2000; 2000US-00173419.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2001-465210/50.
XX
PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT oncogenes and histones, useful for diagnosing and treating, e.g. cancer,
PT autoimmune diseases and infections.
XX
PS Disclosure; Page 4124; 4143pp; English.
XX
CC The present invention relates to oligonucleotides (see AAL26793-AAL34659)
CC encoding polymorphic variants of proteins related to amylases, amyloid
CC proteins, angiopoietin, apoptosis related proteins, cadherin, cyclin,
CC polymerase, oncogenes, histones, kinases, colony stimulating factors,
CC complement related proteins, cytochromes, kinesins, cytokines,

CC interferons, interleukins, G-protein coupled receptors and thioesterases.
CC The present sequence is a peptide encoded by one such oligonucleotide.
CC The oligonucleotides and the peptides encoded by them may be used in the
CC prevention, diagnosis and treatment of diseases associated with
CC inappropriate expression of the proteins listed above. Disorders that may
CC be prevented, diagnosed and/or treated include multifactorial diseases
CC with a genetic component, such as autoimmune diseases (e.g. rheumatoid
CC arthritis, multiple sclerosis, diabetes, systemic lupus erythromatosus
CC and Grave's disease), inflammation, cancer (e.g. cancers of the bladder,
CC brain, breast, colon and kidney, leukaemia), diseases of the nervous
CC system and an infection of pathogenic organisms
XX
SQ Sequence 14 AA;

Query Match 100.0%; Score 29; DB 4; Length 14;
Best Local Similarity 80.0%; Pred. No. 74;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 GPXTW 8
|| ||
Db 7 GPATW 11

RESULT 2
AAM98462
ID AAM98462 standard; peptide; 14 AA.
XX
AC AAM98462;
XX
DT 24-JAN-2002 (first entry)
XX
DE Human peptide #1737 encoded by a SNP oligonucleotide.
XX
KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW complement related protein; cytochrome; kinesin; cytokine; interferon;
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;
KW nervous system disease.
XX
OS Homo sapiens.
XX
PN WO200147944-A2.
XX
PD 05-JUL-2001.
XX
PF 28-DEC-2000; 2000WO-US035498.
XX
PR 28-DEC-1999; 99US-0173419P.
PR 27-DEC-2000; 2000US-00173419.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2001-465210/50.
XX
PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT oncogenes and histones, useful for diagnosing and treating, e.g. cancer,
PT autoimmune diseases and infections.
XX
PS Disclosure; Page 4049; 4143pp; English.
XX
CC The present invention relates to oligonucleotides (see AAL26793-AAL34659)
CC encoding polymorphic variants of proteins related to amylases, amyloid
CC proteins, angiopoietin, apoptosis related proteins, cadherin, cyclin,
CC polymerase, oncogenes, histones, kinases, colony stimulating factors,
CC complement related proteins, cytochromes, kinesins, cytokines,
CC interferons, interleukins, G-protein coupled receptors and thioesterases.
CC The present sequence is a peptide encoded by one such oligonucleotide.
CC The oligonucleotides and the peptides encoded by them may be used in the
CC prevention, diagnosis and treatment of diseases associated with
CC inappropriate expression of the proteins listed above. Disorders that may
CC be prevented, diagnosed and/or treated include multifactorial diseases
CC with a genetic component, such as autoimmune diseases (e.g. rheumatoid
CC arthritis, multiple sclerosis, diabetes, systemic lupus erythromatosus
CC and Grave's disease), inflammation, cancer (e.g. cancers of the bladder,
CC brain, breast, colon and kidney, leukaemia), diseases of the nervous
CC system and an infection of pathogenic organisms
XX
SQ Sequence 14 AA;

Query Match 100.0%; Score 29; DB 4; Length 14;
Best Local Similarity 80.0%; Pred. No. 74;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 GPXTW 8
|| ||
Db 10 GPSTW 14

RESULT 3
ADU91962
ID ADU91962 standard; peptide; 17 AA.
XX
AC ADU91962;
XX
DT 10-FEB-2005 (first entry)
XX
DE EPO-R agonist SEQ ID NO 103.
XX
KW erythropoietin receptor; EPO-R; erythropoietin; renal failure;
KW autoimmune disease; cystic fibrosis; anemia; inflammation;
KW spinal cord injury; aging; neurological disease; nephrotropic;
KW antianemic; immunosuppressive; CNS-Gen.; neuroprotective;
KW respiratory-Gen.; antiinflammatory; vulnerary; nootropic; cytostatic;
KW hemostatic; cyclic.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "Acetylated residue"
FT Disulfide-bond 4. .13
FT Modified-site 17
FT /note= "C-terminal amide"
XX
PN WO2004101611-A2.
XX
PD 25-NOV-2004.
XX
PF 12-MAY-2004; 2004WO-US014886.
XX
PR 12-MAY-2003; 2003US-0470245P.
XX
PA (AFFY-) AFFYMAX INC.
XX
PI Yin K, Holmes C, Lalonde G, Balu P, Schatz PJ, Tumelty D;
XX
DR WPI; 2005-039329/04.
XX
PT New peptide comprising specified sequence of amino acid is erythropoietin
PT receptor agonist useful for treating e.g. anemia, beta-thalassemia, renal
PT disorders.
XX
PS Disclosure; SEQ ID NO 103; 83pp; English.
XX
CC This invention describes a novel peptide which is an erythropoietin
CC receptor (EPO-R) activator. The peptide forms a dimer comprising a
CC linking moiety connecting two peptide chains composed of ADU91861. The N-
CC terminal of the peptide is acetylated. The EPO-R activator further
CC comprises at least one water soluble polymer, preferably polyethylene
CC glycol (PEG) covalently bound to the peptide and a spacer moiety. The
CC products of the invention are used for treating disorders associated with

CC deficiency of erythropoietin or low or defective red blood cell
CC population, end stage renal failure or dialysis, anemia associated with
CC AIDS, autoimmune disease or malignancy, beta-thalassemia, cystic
CC fibrosis, early anemia of prematurity, anemia associated with chronic
CC inflammatory disease, spinal cord injury, acute blood loss, aging and
CC neoplastic disease states accompanied by abnormal erythropoiesis. The
CC peptide compounds are potent agonists of erythropoietin receptor and have
CC nephrotropic, antianemic, immunosuppressive, CNS-Gen., neuroprotective,
CC respiratory-Gen., antiinflammatory, vulnerary, nootropic, cytostatic and
CC hemostatic activity. This sequence represents a peptide which acts as an
CC erythropoietin receptor (EPO-R) agonist.
XX
SQ Sequence 17 AA;

Query Match 100.0%; Score 29; DB 9; Length 17;
Best Local Similarity 80.0%; Pred. No. 90;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 GPXTW 8
|| ||
Db 7 GPTTW 11

RESULT 4
ADY54197
ID ADY54197 standard; peptide; 17 AA.
XX
AC ADY54197;
XX
DT 19-MAY-2005 (first entry)
XX
DE Amino acid sequence of mutated EMP-1 #4.
XX
KW cytostatic; anti-HIV; hypotensive; neuroprotective; cardiovascular-Gen.;
KW nootropic; hepatotropic; virucide; antiinflammatory; immunosuppressive;
KW antiallergic; antimicrobial; neuroleptic; gynecological; anorectic;
KW antiarteriosclerotic; gastrointestinal-Gen.; endocrine-Gen; neoplasm;
KW hematological disease; erythropoietin peptide mimetic; EPM;
KW EPO mimetic peptide-1; EMP-1; multiple sclerosis; brain tumor; cancer;
KW hepatitis; anemia; pregnancy; menstrual disorder; rheumatoid arthritis;
KW AIDS; viral disease; metabolic disease; autoimmune disease;
KW inflammatory disease; allergy; microbial infection;
KW cardiovascular disease; genetic disease; neurodegenerative disease;
KW hematopoietic cell disorder; endocrine disorder;
KW gastrointestinal disease; hypertension; arterial sclerosis.
XX
OS Synthetic.
XX
PN WO2005021579-A2.
XX
PD 10-MAR-2005.
XX
PF 30-AUG-2004; 2004WO-US027949.
XX
PR 28-AUG-2003; 2003WO-US026818.
PR 10-MAR-2004; 2004US-0551552P.
XX
PA (BIOR-) BIOREXIS PHARM CORP.
XX
PI Sadeghi H, Turner AJ;
XX
DR WPI; 2005-214540/22.
XX
PT Novel erythropoietin (EPO) peptide mimetic, having first modification of
PT cysteine residue of EPO mimetic peptides (EMP)-1, to reduces disulfide
PT bond formation, and second modification such that peptide exhibits EMP-1
PT activity.
XX
PS Example 2; SEQ ID NO 51; 158pp; English.
XX
CC The specification describes an erythropoietin (EPO) peptide mimetic
CC (EPM), comprising a modification of at least one cysteine residue of EPO
CC mimetic peptide (EMP)-1 that substantially reduces disulfide bond
CC formation, and a second modification such that the peptide exhibits EMP-1
CC activity. The first modification comprises the deletion or substitution
CC of at least one cysteine residue in EMP-1, and the second modification
CC comprises the addition of a linker group that is covalently bonded to the
CC C-terminal amino acid or N-terminal amino acid of EMP-1. EPM peptides of
CC the invention are useful for treating or preventing diseases, such as
CC multiple sclerosis, brain tumor, skin cancer, hepatitis B, hepatitis C,
CC anemia, beta-thalassemia, pregnancy or menstrual disorders, rheumatoid
CC arthritis, AIDS, cancer, viral disease, metabolic disease, obesity,
CC autoimmune disease, inflammatory disease, allergy, graft-versus-host
CC disease, systemic microbial infection, cardiovascular disease, psychosis,
CC genetic diseases, neurodegenerative diseases, disorders of hematopoietic
CC cells, diseases of the endocrine system or reproductive systems,
CC gastrointestinal diseases, diabetes, asthma, or HIV infections,
CC hypertension, hypercholesterolemia, arterial sclerosis, arthritis or
CC Alzheimer's disease. The present sequence represents a mutated EMP-1,
CC used to produce an EPM of the invention.
XX
SQ Sequence 17 AA;

Query Match 100.0%; Score 29; DB 9; Length 17;
Best Local Similarity 80.0%; Pred. No. 90;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 GPXTW 8
|| ||
Db 9 GPTTW 13

RESULT 5
AAY13704
ID AAY13704 standard; peptide; 20 AA.
XX
AC AAY13704;
XX
DT 06-SEP-1999 (first entry)
XX
DE Erythropoietin receptor (EPO-R) binding peptide.
XX
KW Erythropoietin; EPO; receptor; EPO deficiency; renal failure; AIDS;
KW dialysis; anaemia; autoimmune disease; chronic inflammatory disease;
KW malignancy; beta-thalassaemia; cystic fibrosis; prematurity; blood loss;
KW spinal cord injury; aging; neoplastic disease; erythropoiesis; EPO-R.
XX
OS Synthetic.
XX
PN WO9640749-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US009810.
XX
PR 07-JUN-1995; 95US-00484631.
PR 07-JUN-1995; 95US-00484635.
XX
PA (JOHJ) JOHNSON & JOHNSON CORP.
PA (AFFY-) AFFYMAX TECHNOLOGIES NV.
XX
PI Wrighton NC, Dower WJ, Chang RS, Kashyap AK, Jolliffe LK;
PI Johnson D, Mulcahy L;
XX
DR WPI; 1997-052225/05.
XX
PT Erythropoietin receptor binding peptide - useful for treating disorders
PT characterised by deficiency of EPO, or low or defective red blood cell
PT population.
XX
PS Disclosure; Fig 2; 95pp; English.
XX
CC The invention describes a peptide of 10-40 amino acid residues which
CC binds to erythropoietin (EPO) receptor and which includes the amino acid
CC sequence Cys-Xaa1-Xaa2-Gly-Pro-Xaa3-Thr-Trp-Xaa4-Cys, where Xaa1 = Arg,
CC His, Leu or Trp, Xaa2 = Met, Phe or Ile, Xaa3 = 1 of the 20 genetically

CC coded L-amino acids, and Xaa4 = Asp, Glu, Ile, Leu or Val. Optionally,
CC the peptide may be cyclised or dimerised. The peptide can be used to
CC treat a patient having a disorder characterised by a deficiency of EPO or
CC a low or defective red blood cell population. It can be used to treat end
CC stage renal failure or dialysis; anaemia associated with AIDS, autoimmune
CC disease, chronic inflammatory diseases or malignancy; beta-thalassaemia;
CC cystic fibrosis; early anaemia of prematurity; spinal cord injury; acute
CC blood loss; aging; and neoplastic disease states accompanied by abnormal
CC erythropoiesis. The peptides can also be used as reagents for detecting
CC EPO receptors on living cells, in biological fluids, in tissue
CC homogenates, etc. Sequences AAY13662-735 are representative peptides of
CC the invention
XX
SQ Sequence 20 AA;

Query Match 100.0%; Score 29; DB 2; Length 20;
Best Local Similarity 80.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 GPXTW 8
|| ||
Db 9 GPTTW 13

RESULT 6
AAY13696
ID AAY13696 standard; peptide; 20 AA.
XX
AC AAY13696;
XX
DT 06-SEP-1999 (first entry)
XX
DE Erythropoietin receptor (EPO-R) binding peptide.
XX
KW Erythropoietin; EPO; receptor; EPO deficiency; renal failure; AIDS;
KW dialysis; anaemia; autoimmune disease; chronic inflammatory disease;
KW malignancy; beta-thalassaemia; cystic fibrosis; prematurity; blood loss;
KW spinal cord injury; aging; neoplastic disease; erythropoiesis; EPO-R.
XX
OS Synthetic.
XX
PN WO9640749-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US009810.
XX
PR 07-JUN-1995; 95US-00484631.
PR 07-JUN-1995; 95US-00484635.
XX
PA (JOHJ) JOHNSON & JOHNSON CORP.
PA (AFFY-) AFFYMAX TECHNOLOGIES NV.
XX
PI Wrighton NC, Dower WJ, Chang RS, Kashyap AK, Jolliffe LK;
PI Johnson D, Mulcahy L;
XX
DR WPI; 1997-052225/05.
XX
PT Erythropoietin receptor binding peptide - useful for treating disorders
PT characterised by deficiency of EPO, or low or defective red blood cell
PT population.
XX
PS Disclosure; Fig 2; 95pp; English.
XX
CC The invention describes a peptide of 10-40 amino acid residues which
CC binds to erythropoietin (EPO) receptor and which includes the amino acid
CC sequence Cys-Xaa1-Xaa2-Gly-Pro-Xaa3-Thr-Trp-Xaa4-Cys, where Xaa1 = Arg,
CC His, Leu or Trp, Xaa2 = Met, Phe or Ile, Xaa3 = 1 of the 20 genetically
CC coded L-amino acids, and Xaa4 = Asp, Glu, Ile, Leu or Val. Optionally,
CC the peptide may be cyclised or dimerised. The peptide can be used to
CC treat a patient having a disorder characterised by a deficiency of EPO or
CC a low or defective red blood cell population. It can be used to treat end
CC stage renal failure or dialysis; anaemia associated with AIDS, autoimmune
CC disease, chronic inflammatory diseases or malignancy; beta-thalassaemia;
CC cystic fibrosis; early anaemia of prematurity; spinal cord injury; acute
CC blood loss; aging; and neoplastic disease states accompanied by abnormal
CC erythropoiesis. The peptides can also be used as reagents for detecting
CC EPO receptors on living cells, in biological fluids, in tissue
CC homogenates, etc. Sequences AAY13662-735 are representative peptides of
CC the invention
XX
SQ Sequence 20 AA;

Query Match 100.0%; Score 29; DB 2; Length 20;
Best Local Similarity 80.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 GPXTW 8
|| ||
Db 9 GPTTW 13

RESULT 7
AAY13650
ID AAY13650 standard; peptide; 20 AA.
XX
AC AAY13650;
XX
DT 06-SEP-1999 (first entry)
XX
DE Erythropoietin receptor (EPO-R) binding peptide.
XX
KW Erythropoietin; EPO; receptor; EPO deficiency; renal failure; AIDS;
KW dialysis; anaemia; autoimmune disease; chronic inflammatory disease;
KW malignancy; beta-thalassaemia; cystic fibrosis; prematurity; blood loss;
KW spinal cord injury; aging; neoplastic disease; erythropoiesis; EPO-R.
XX
OS Synthetic.
XX
PN WO9640749-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US009810.
XX
PR 07-JUN-1995; 95US-00484631.
PR 07-JUN-1995; 95US-00484635.
XX
PA (JOHJ) JOHNSON & JOHNSON CORP.
PA (AFFY-) AFFYMAX TECHNOLOGIES NV.
XX
PI Wrighton NC, Dower WJ, Chang RS, Kashyap AK, Jolliffe LK;
PI Johnson D, Mulcahy L;
XX
DR WPI; 1997-052225/05.
XX
PT Erythropoietin receptor binding peptide - useful for treating disorders
PT characterised by deficiency of EPO, or low or defective red blood cell
PT population.
XX
PS Claim 6; Page 68; 95pp; English.
XX
CC The invention describes a peptide of 10-40 amino acid residues which
CC binds to erythropoietin (EPO) receptor and which includes the amino acid
CC sequence Cys-Xaa1-Xaa2-Gly-Pro-Xaa3-Thr-Trp-Xaa4-Cys, where Xaa1 = Arg,
CC His, Leu or Trp, Xaa2 = Met, Phe or Ile, Xaa3 = 1 of the 20 genetically
CC coded L-amino acids, and Xaa4 = Asp, Glu, Ile, Leu or Val. Optionally,
CC the peptide may be cyclised or dimerised. The peptide can be used to
CC treat a patient having a disorder characterised by a deficiency of EPO or
CC a low or defective red blood cell population. It can be used to treat end
CC stage renal failure or dialysis; anaemia associated with AIDS, autoimmune
CC disease, chronic inflammatory diseases or malignancy; beta-thalassaemia;
CC cystic fibrosis; early anaemia of prematurity; spinal cord injury; acute
CC blood loss; aging; and neoplastic disease states accompanied by abnormal
CC erythropoiesis. The peptides can also be used as reagents for detecting
CC EPO receptors on living cells, in biological fluids, in tissue

```
CC   homogenates, etc. Sequences AAY13624-661 represent specifc examples of
CC   EPO-R binding peptides
XX
SQ   Sequence 20 AA;

  Query Match             100.0%;  Score 29;  DB 2;  Length 20;
  Best Local Similarity   80.0%;  Pred. No. 1.1e+02;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 GPXTW 8
              || ||
Db          9 GPATW 13


RESULT 8
AAY13728
ID   AAY13728 standard; peptide; 20 AA.
XX
AC   AAY13728;
XX
DT   06-SEP-1999  (first entry)
XX
DE   Erythropoietin receptor (EPO-R) binding peptide.
XX
KW   Erythropoietin; EPO; receptor; EPO deficiency; renal failure; AIDS;
KW   dialysis; anaemia; autoimmune disease; chronic inflammatory disease;
KW   malignancy; beta-thalassaemia; cystic fibrosis; prematurity; blood loss;
KW   spinal cord injury; aging; neoplastic disease; erythropoiesis; EPO-R.
XX
OS   Synthetic.
XX
PN   WO9640749-A1.
XX
PD   19-DEC-1996.
XX
PF   07-JUN-1996;   96WO-US009810.
XX
PR   07-JUN-1995;   95US-00484631.
PR   07-JUN-1995;   95US-00484635.
XX
PA   (JOHJ ) JOHNSON & JOHNSON CORP.
PA   (AFFY-) AFFYMAX TECHNOLOGIES NV.
XX
PI   Wrighton NC,  Dower WJ,  Chang RS,  Kashyap AK,  Jolliffe LK;
PI   Johnson D,  Mulcahy L;
XX
DR   WPI; 1997-052225/05.
XX
PT   Erythropoietin receptor binding peptide - useful for treating disorders
PT   characterised by deficiency of EPO, or low or defective red blood cell
PT   population.
XX
PS   Disclosure; Fig 2; 95pp; English.
XX
CC   The invention describes a peptide of 10-40 amino acid residues which
CC   binds to erythropoietin (EPO) receptor and which includes the amino acid
CC   sequence Cys-Xaa1-Xaa2-Gly-Pro-Xaa3-Thr-Trp-Xaa4-Cys, where Xaa1 = Arg,
CC   His, Leu or Trp, Xaa2 = Met, Phe or Ile, Xaa3 = 1 of the 20 genetically
CC   coded L-amino acids, and Xaa4 = Asp, Glu, Ile, Leu or Val. Optionally,
CC   the peptide may be cyclised or dimerised. The peptide can be used to
CC   treat a patient having a disorder characterised by a deficiency of EPO or
CC   a low or defective red blood cell population. It can be used to treat end
CC   stage renal failure or dialysis; anaemia associated with AIDS, autoimmune
CC   disease, chronic inflammatory diseases or malignancy; beta-thalassaemia;
CC   cystic fibrosis; early anaemia of prematurity; spinal cord injury; acute
CC   blood loss; aging; and neoplastic disease states accompanied by abnormal
CC   erythropoiesis. The peptides can also be used as reagents for detecting
CC   EPO receptors on living cells, in biological fluids, in tissue
CC   homogenates, etc. Sequences AAY13662-735 are representative peptides of
CC   the invention
XX
SQ   Sequence 20 AA;

  Query Match             100.0%;  Score 29;  DB 2;  Length 20;
  Best Local Similarity   80.0%;  Pred. No. 1.1e+02;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 GPXTW 8
              || ||
Db          9 GPTTW 13


RESULT 9
AAY13688
ID   AAY13688 standard; peptide; 20 AA.
XX
AC   AAY13688;
XX
DT   06-SEP-1999  (first entry)
XX
DE   Erythropoietin receptor (EPO-R) binding peptide.
XX
KW   Erythropoietin; EPO; receptor; EPO deficiency; renal failure; AIDS;
KW   dialysis; anaemia; autoimmune disease; chronic inflammatory disease;
KW   malignancy; beta-thalassaemia; cystic fibrosis; prematurity; blood loss;
KW   spinal cord injury; aging; neoplastic disease; erythropoiesis; EPO-R.
XX
OS   Synthetic.
XX
PN   WO9640749-A1.
XX
PD   19-DEC-1996.
XX
PF   07-JUN-1996;   96WO-US009810.
XX
PR   07-JUN-1995;   95US-00484631.
PR   07-JUN-1995;   95US-00484635.
XX
PA   (JOHJ ) JOHNSON & JOHNSON CORP.
PA   (AFFY-) AFFYMAX TECHNOLOGIES NV.
XX
PI   Wrighton NC,  Dower WJ,  Chang RS,  Kashyap AK,  Jolliffe LK;
PI   Johnson D,  Mulcahy L;
XX
DR   WPI; 1997-052225/05.
XX
PT   Erythropoietin receptor binding peptide - useful for treating disorders
PT   characterised by deficiency of EPO, or low or defective red blood cell
PT   population.
XX
PS   Disclosure; Fig 2; 95pp; English.
XX
CC   The invention describes a peptide of 10-40 amino acid residues which
CC   binds to erythropoietin (EPO) receptor and which includes the amino acid
CC   sequence Cys-Xaa1-Xaa2-Gly-Pro-Xaa3-Thr-Trp-Xaa4-Cys, where Xaa1 = Arg,
CC   His, Leu or Trp, Xaa2 = Met, Phe or Ile, Xaa3 = 1 of the 20 genetically
CC   coded L-amino acids, and Xaa4 = Asp, Glu, Ile, Leu or Val. Optionally,
CC   the peptide may be cyclised or dimerised. The peptide can be used to
CC   treat a patient having a disorder characterised by a deficiency of EPO or
CC   a low or defective red blood cell population. It can be used to treat end
CC   stage renal failure or dialysis; anaemia associated with AIDS, autoimmune
CC   disease, chronic inflammatory diseases or malignancy; beta-thalassaemia;
CC   cystic fibrosis; early anaemia of prematurity; spinal cord injury; acute
CC   blood loss; aging; and neoplastic disease states accompanied by abnormal
CC   erythropoiesis. The peptides can also be used as reagents for detecting
CC   EPO receptors on living cells, in biological fluids, in tissue
CC   homogenates, etc. Sequences AAY13662-735 are representative peptides of
CC   the invention
XX
SQ   Sequence 20 AA;

  Query Match             100.0%;  Score 29;  DB 2;  Length 20;
  Best Local Similarity   80.0%;  Pred. No. 1.1e+02;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 GPXTW 8
```

```
              || ||
Db          9 GPSTW 13


RESULT 10
AAY13687
ID   AAY13687 standard; peptide; 20 AA.
XX
AC   AAY13687;
XX
DT   06-SEP-1999  (first entry)
XX
DE   Erythropoietin receptor (EPO-R) binding peptide.
XX
KW   Erythropoietin; EPO; receptor; EPO deficiency; renal failure; AIDS;
KW   dialysis; anaemia; autoimmune disease; chronic inflammatory disease;
KW   malignancy; beta-thalassaemia; cystic fibrosis; prematurity; blood loss;
KW   spinal cord injury; aging; neoplastic disease; erythropoiesis; EPO-R.
XX
OS   Synthetic.
XX
PN   WO9640749-A1.
XX
PD   19-DEC-1996.
XX
PF   07-JUN-1996;   96WO-US009810.
XX
PR   07-JUN-1995;   95US-00484631.
PR   07-JUN-1995;   95US-00484635.
XX
PA   (JOHJ ) JOHNSON & JOHNSON CORP.
PA   (AFFY-) AFFYMAX TECHNOLOGIES NV.
XX
PI   Wrighton NC,  Dower WJ,  Chang RS,  Kashyap AK,  Jolliffe LK;
PI   Johnson D,  Mulcahy L;
XX
DR   WPI; 1997-052225/05.
XX
PT   Erythropoietin receptor binding peptide - useful for treating disorders
PT   characterised by deficiency of EPO, or low or defective red blood cell
PT   population.
XX
PS   Disclosure; Fig 2; 95pp; English.
XX
CC   The invention describes a peptide of 10-40 amino acid residues which
CC   binds to erythropoietin (EPO) receptor and which includes the amino acid
CC   sequence Cys-Xaa1-Xaa2-Gly-Pro-Xaa3-Thr-Trp-Xaa4-Cys, where Xaa1 = Arg,
CC   His, Leu or Trp, Xaa2 = Met, Phe or Ile, Xaa3 = 1 of the 20 genetically
CC   coded L-amino acids, and Xaa4 = Asp, Glu, Ile, Leu or Val. Optionally,
CC   the peptide may be cyclised or dimerised. The peptide can be used to
CC   treat a patient having a disorder characterised by a deficiency of EPO or
CC   a low or defective red blood cell population. It can be used to treat end
CC   stage renal failure or dialysis; anaemia associated with AIDS, autoimmune
CC   disease, chronic inflammatory diseases or malignancy; beta-thalassaemia;
CC   cystic fibrosis; early anaemia of prematurity; spinal cord injury; acute
CC   blood loss; aging; and neoplastic disease states accompanied by abnormal
CC   erythropoiesis. The peptides can also be used as reagents for detecting
CC   EPO receptors on living cells, in biological fluids, in tissue
CC   homogenates, etc. Sequences AAY13662-735 are representative peptides of
CC   the invention
XX
SQ   Sequence 20 AA;

  Query Match             100.0%;  Score 29;  DB 2;  Length 20;
  Best Local Similarity   80.0%;  Pred. No. 1.1e+02;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 GPXTW 8
              || ||
Db          9 GPTTW 13


RESULT 11
```

```
AAY13705
ID   AAY13705 standard; peptide; 20 AA.
XX
AC   AAY13705;
XX
DT   06-SEP-1999  (first entry)
XX
DE   Erythropoietin receptor (EPO-R) binding peptide.
XX
KW   Erythropoietin; EPO; receptor; EPO deficiency; renal failure; AIDS;
KW   dialysis; anaemia; autoimmune disease; chronic inflammatory disease;
KW   malignancy; beta-thalassaemia; cystic fibrosis; prematurity; blood loss;
KW   spinal cord injury; aging; neoplastic disease; erythropoiesis; EPO-R.
XX
OS   Synthetic.
XX
PN   WO9640749-A1.
XX
PD   19-DEC-1996.
XX
PF   07-JUN-1996;   96WO-US009810.
XX
PR   07-JUN-1995;   95US-00484631.
PR   07-JUN-1995;   95US-00484635.
XX
PA   (JOHJ ) JOHNSON & JOHNSON CORP.
PA   (AFFY-) AFFYMAX TECHNOLOGIES NV.
XX
PI   Wrighton NC,  Dower WJ,  Chang RS,  Kashyap AK,  Jolliffe LK;
PI   Johnson D,  Mulcahy L;
XX
DR   WPI; 1997-052225/05.
XX
PT   Erythropoietin receptor binding peptide - useful for treating disorders
PT   characterised by deficiency of EPO, or low or defective red blood cell
PT   population.
XX
PS   Disclosure; Fig 2; 95pp; English.
XX
CC   The invention describes a peptide of 10-40 amino acid residues which
CC   binds to erythropoietin (EPO) receptor and which includes the amino acid
CC   sequence Cys-Xaa1-Xaa2-Gly-Pro-Xaa3-Thr-Trp-Xaa4-Cys, where Xaa1 = Arg,
CC   His, Leu or Trp, Xaa2 = Met, Phe or Ile, Xaa3 = 1 of the 20 genetically
CC   coded L-amino acids, and Xaa4 = Asp, Glu, Ile, Leu or Val. Optionally,
CC   the peptide may be cyclised or dimerised. The peptide can be used to
CC   treat a patient having a disorder characterised by a deficiency of EPO or
CC   a low or defective red blood cell population. It can be used to treat end
CC   stage renal failure or dialysis; anaemia associated with AIDS, autoimmune
CC   disease, chronic inflammatory diseases or malignancy; beta-thalassaemia;
CC   cystic fibrosis; early anaemia of prematurity; spinal cord injury; acute
CC   blood loss; aging; and neoplastic disease states accompanied by abnormal
CC   erythropoiesis. The peptides can also be used as reagents for detecting
CC   EPO receptors on living cells, in biological fluids, in tissue
CC   homogenates, etc. Sequences AAY13662-735 are representative peptides of
CC   the invention
XX
SQ   Sequence 20 AA;

  Query Match             100.0%;  Score 29;  DB 2;  Length 20;
  Best Local Similarity   80.0%;  Pred. No. 1.1e+02;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 GPXTW 8
              || ||
Db          9 GPTTW 13


RESULT 12
AAY26368
ID   AAY26368 standard; peptide; 20 AA.
XX
AC   AAY26368;
XX
```

```
DT   06-SEP-1999  (first entry)
XX
DE   Erythropoietin receptor (EPO-R) binding peptide.
XX
KW   Erythropoietin; EPO; receptor; EPO deficiency; renal failure; AIDS;
KW   dialysis; anaemia; autoimmune disease; chronic inflammatory disease;
KW   malignancy; beta-thalassaemia; cystic fibrosis; prematurity; blood loss;
KW   spinal cord injury; aging; neoplastic disease; erythropoiesis; EPO-R.
XX
OS   Synthetic.
XX
PN   WO9640749-A1.
XX
PD   19-DEC-1996.
XX
PF   07-JUN-1996;   96WO-US009810.
XX
PR   07-JUN-1995;   95US-00484631.
PR   07-JUN-1995;   95US-00484635.
XX
PA   (JOHJ ) JOHNSON & JOHNSON CORP.
PA   (AFFY-) AFFYMAX TECHNOLOGIES NV.
XX
PI   Wrighton NC,  Dower WJ,  Chang RS,  Kashyap AK,  Jolliffe LK;
PI   Johnson D,  Mulcahy L;
XX
DR   WPI; 1997-052225/05.
XX
PT   Erythropoietin receptor binding peptide - useful for treating disorders
PT   characterised by deficiency of EPO, or low or defective red blood cell
PT   population.
XX
PS   Disclosure; Page 16; 95pp; English.
XX
CC   The invention describes a peptide of 10-40 amino acid residues which
CC   binds to erythropoietin (EPO) receptor and which includes the amino acid
CC   sequence Cys-Xaa1-Xaa2-Gly-Pro-Xaa3-Thr-Trp-Xaa4-Cys, where Xaa1 = Arg,
CC   His, Leu or Trp, Xaa2 = Met, Phe or Ile, Xaa3 = 1 of the 20 genetically
CC   coded L-amino acids, and Xaa4 = Asp, Glu, Ile, Leu or Val. Optionally,
CC   the peptide may be cyclised or dimerised. The peptide can be used to
CC   treat a patient having a disorder characterised by a deficiency of EPO or
CC   a low or defective red blood cell population. It can be used to treat end
CC   stage renal failure or dialysis; anaemia associated with AIDS, autoimmune
CC   disease, chronic inflammatory diseases or malignancy; beta-thalassaemia;
CC   cystic fibrosis; early anaemia of prematurity; spinal cord injury; acute
CC   blood loss; aging; and neoplastic disease states accompanied by abnormal
CC   erythropoiesis. The peptides can also be used as reagents for detecting
CC   EPO receptors on living cells, in biological fluids, in tissue
CC   homogenates, etc. Sequences AAY26352-548 are representative peptides
CC   falling within the above peptide motif and isolated during the affinity
CC   selection process
XX
SQ   Sequence 20 AA;

  Query Match             100.0%;  Score 29;  DB 2;  Length 20;
  Best Local Similarity   80.0%;  Pred. No. 1.1e+02;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 GPXTW 8
              || ||
Db          9 GPTTW 13


RESULT 13
AAY13672
ID   AAY13672 standard; peptide; 20 AA.
XX
AC   AAY13672;
XX
DT   06-SEP-1999  (first entry)
XX
DE   Erythropoietin receptor (EPO-R) binding peptide.
XX
KW   Erythropoietin; EPO; receptor; EPO deficiency; renal failure; AIDS;
KW   dialysis; anaemia; autoimmune disease; chronic inflammatory disease;
KW   malignancy; beta-thalassaemia; cystic fibrosis; prematurity; blood loss;
KW   spinal cord injury; aging; neoplastic disease; erythropoiesis; EPO-R.
XX
OS   Synthetic.
XX
PN   WO9640749-A1.
XX
PD   19-DEC-1996.
XX
PF   07-JUN-1996;   96WO-US009810.
XX
PR   07-JUN-1995;   95US-00484631.
PR   07-JUN-1995;   95US-00484635.
XX
PA   (JOHJ ) JOHNSON & JOHNSON CORP.
PA   (AFFY-) AFFYMAX TECHNOLOGIES NV.
XX
PI   Wrighton NC,  Dower WJ,  Chang RS,  Kashyap AK,  Jolliffe LK;
PI   Johnson D,  Mulcahy L;
XX
DR   WPI; 1997-052225/05.
XX
PT   Erythropoietin receptor binding peptide - useful for treating disorders
PT   characterised by deficiency of EPO, or low or defective red blood cell
PT   population.
XX
PS   Disclosure; Fig 2; 95pp; English.
XX
CC   The invention describes a peptide of 10-40 amino acid residues which
CC   binds to erythropoietin (EPO) receptor and which includes the amino acid
CC   sequence Cys-Xaa1-Xaa2-Gly-Pro-Xaa3-Thr-Trp-Xaa4-Cys, where Xaa1 = Arg,
CC   His, Leu or Trp, Xaa2 = Met, Phe or Ile, Xaa3 = 1 of the 20 genetically
CC   coded L-amino acids, and Xaa4 = Asp, Glu, Ile, Leu or Val. Optionally,
CC   the peptide may be cyclised or dimerised. The peptide can be used to
CC   treat a patient having a disorder characterised by a deficiency of EPO or
CC   a low or defective red blood cell population. It can be used to treat end
CC   stage renal failure or dialysis; anaemia associated with AIDS, autoimmune
CC   disease, chronic inflammatory diseases or malignancy; beta-thalassaemia;
CC   cystic fibrosis; early anaemia of prematurity; spinal cord injury; acute
CC   blood loss; aging; and neoplastic disease states accompanied by abnormal
CC   erythropoiesis. The peptides can also be used as reagents for detecting
CC   EPO receptors on living cells, in biological fluids, in tissue
CC   homogenates, etc. Sequences AAY13662-735 are representative peptides of
CC   the invention
XX
SQ   Sequence 20 AA;

  Query Match             100.0%;  Score 29;  DB 2;  Length 20;
  Best Local Similarity   80.0%;  Pred. No. 1.1e+02;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 GPXTW 8
              || ||
Db          9 GPATW 13


RESULT 14
AAY13706
ID   AAY13706 standard; peptide; 20 AA.
XX
AC   AAY13706;
XX
DT   06-SEP-1999  (first entry)
XX
DE   Erythropoietin receptor (EPO-R) binding peptide.
XX
KW   Erythropoietin; EPO; receptor; EPO deficiency; renal failure; AIDS;
KW   dialysis; anaemia; autoimmune disease; chronic inflammatory disease;
KW   malignancy; beta-thalassaemia; cystic fibrosis; prematurity; blood loss;
KW   spinal cord injury; aging; neoplastic disease; erythropoiesis; EPO-R.
XX
```

OS Synthetic.
XX
PN WO9640749-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US009810.
XX
PR 07-JUN-1995; 95US-00484631.
PR 07-JUN-1995; 95US-00484635.
XX
PA (JOHJ) JOHNSON & JOHNSON CORP.
PA (AFFY-) AFFYMAX TECHNOLOGIES NV.
XX
PI Wrighton NC, Dower WJ, Chang RS, Kashyap AK, Jolliffe LK;
PI Johnson D, Mulcahy L;
XX
DR WPI; 1997-052225/05.
XX
PT Erythropoietin receptor binding peptide - useful for treating disorders
PT characterised by deficiency of EPO, or low or defective red blood cell
PT population.
XX
PS Disclosure; Fig 2; 95pp; English.
XX
CC The invention describes a peptide of 10-40 amino acid residues which
CC binds to erythropoietin (EPO) receptor and which includes the amino acid
CC sequence Cys-Xaa1-Xaa2-Gly-Pro-Xaa3-Thr-Trp-Xaa4-Cys, where Xaa1 = Arg,
CC His, Leu or Trp, Xaa2 = Met, Phe or Ile, Xaa3 = 1 of the 20 genetically
CC coded L-amino acids, and Xaa4 = Asp, Glu, Ile, Leu or Val. Optionally,
CC the peptide may be cyclised or dimerised. The peptide can be used to
CC treat a patient having a disorder characterised by a deficiency of EPO or
CC a low or defective red blood cell population. It can be used to treat end
CC stage renal failure or dialysis; anaemia associated with AIDS, autoimmune
CC disease, chronic inflammatory diseases or malignancy; beta-thalassaemia;
CC cystic fibrosis; early anaemia of prematurity; spinal cord injury; acute
CC blood loss; aging; and neoplastic disease states accompanied by abnormal
CC erythropoiesis. The peptides can also be used as reagents for detecting
CC EPO receptors on living cells, in biological fluids, in tissue
CC homogenates, etc. Sequences AAY13662-735 are representative peptides of
CC the invention
XX
SQ Sequence 20 AA;

Query Match 100.0%; Score 29; DB 2; Length 20;
Best Local Similarity 80.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 GPXTW 8
|| ||
Db 9 GPTTW 13

RESULT 15
AAY13679
ID AAY13679 standard; peptide; 20 AA.
XX
AC AAY13679;
XX
DT 06-SEP-1999 (first entry)
XX
DE Erythropoietin receptor (EPO-R) binding peptide.
XX
KW Erythropoietin; EPO; receptor; EPO deficiency; renal failure; AIDS;
KW dialysis; anaemia; autoimmune disease; chronic inflammatory disease;
KW malignancy; beta-thalassaemia; cystic fibrosis; prematurity; blood loss;
KW spinal cord injury; aging; neoplastic disease; erythropoiesis; EPO-R.
XX
OS Synthetic.
XX
PN WO9640749-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US009810.
XX
PR 07-JUN-1995; 95US-00484631.
PR 07-JUN-1995; 95US-00484635.
XX
PA (JOHJ) JOHNSON & JOHNSON CORP.
PA (AFFY-) AFFYMAX TECHNOLOGIES NV.
XX
PI Wrighton NC, Dower WJ, Chang RS, Kashyap AK, Jolliffe LK;
PI Johnson D, Mulcahy L;
XX
DR WPI; 1997-052225/05.
XX
PT Erythropoietin receptor binding peptide - useful for treating disorders
PT characterised by deficiency of EPO, or low or defective red blood cell
PT population.
XX
PS Disclosure; Fig 2; 95pp; English.
XX
CC The invention describes a peptide of 10-40 amino acid residues which
CC binds to erythropoietin (EPO) receptor and which includes the amino acid
CC sequence Cys-Xaa1-Xaa2-Gly-Pro-Xaa3-Thr-Trp-Xaa4-Cys, where Xaa1 = Arg,
CC His, Leu or Trp, Xaa2 = Met, Phe or Ile, Xaa3 = 1 of the 20 genetically
CC coded L-amino acids, and Xaa4 = Asp, Glu, Ile, Leu or Val. Optionally,
CC the peptide may be cyclised or dimerised. The peptide can be used to
CC treat a patient having a disorder characterised by a deficiency of EPO or
CC a low or defective red blood cell population. It can be used to treat end
CC stage renal failure or dialysis; anaemia associated with AIDS, autoimmune
CC disease, chronic inflammatory diseases or malignancy; beta-thalassaemia;
CC cystic fibrosis; early anaemia of prematurity; spinal cord injury; acute
CC blood loss; aging; and neoplastic disease states accompanied by abnormal
CC erythropoiesis. The peptides can also be used as reagents for detecting
CC EPO receptors on living cells, in biological fluids, in tissue
CC homogenates, etc. Sequences AAY13662-735 are representative peptides of
CC the invention
XX
SQ Sequence 20 AA;

Query Match 100.0%; Score 29; DB 2; Length 20;
Best Local Similarity 80.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 GPXTW 8
|| ||
Db 9 GPATW 13

Search completed: March 31, 2006, 16:22:27
Job time : 39.5572 secs

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: March 31, 2006, 16:22:51 ; Search time 6.21891 Seconds
(without alignments)
154.717 Million cell updates/sec

Title: US-10-609-217-124
Perfect score: 29
Sequence: 1 XXXGPXTWXX 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	29	100.0	227	2	C39925	hypothetical prote
2	29	100.0	237	2	G87286	conserved hypothet
3	29	100.0	246	2	B86784	hypothetical prote
4	29	100.0	263	2	T48742	hypothetical prote
5	29	100.0	279	2	AC3647	cellobiose phospho
6	29	100.0	332	2	B87356	sugar ABC transpor
7	29	100.0	334	2	JC6082	proximal sequence
8	29	100.0	371	2	T42623	probable sugar tra
9	29	100.0	433	2	T44587	cytochrome P450 ho
10	29	100.0	433	2	S63383	POP2 protein - yea
11	29	100.0	463	2	S36507	L2 protein - human
12	29	100.0	464	2	S36582	L2 protein - human
13	29	100.0	475	2	H84137	hypothetical prote
14	29	100.0	505	1	D70703	tldD homolog Rv231
15	29	100.0	506	2	D90207	conserved hypothet
16	29	100.0	536	2	D83622	arylsulfatase PA01
17	29	100.0	537	1	FOMVM7	gag polyprotein -
18	29	100.0	546	2	T40888	probable glucose t
19	29	100.0	565	2	T14732	probable beta-gluc
20	29	100.0	781	2	T49472	hormone-sensitive
21	29	100.0	821	2	B84509	probable Na/H anti
22	29	100.0	852	2	A34373	histidine-rich cal
23	29	100.0	1208	2	T00362	hypothetical prote
24	28	96.6	19	1	EWSMAN	ancovenin - Strept
25	28	96.6	60	2	S78724	protein YKL033w-a
26	28	96.6	62	2	E84394	hypothetical prote
27	28	96.6	68	2	S36976	hypothetical prote
28	28	96.6	68	2	B43940	lactococcin B prec
29	28	96.6	93	2	T06470	probable chitinase
30	28	96.6	99	2	D75378	hypothetical prote
31	28	96.6	118	2	T17205	hypothetical prote
32	28	96.6	118	2	S59930	hypothetical prote
33	28	96.6	149	2	T26485	hypothetical prote
34	28	96.6	164	2	T04299	pathogenesis-relat
35	28	96.6	164	2	F83798	hypothetical prote
36	28	96.6	167	2	S51679	pathogenesis-relat
37	28	96.6	167	2	S14969	pathogenesis-relat
38	28	96.6	175	2	S43894	pathogenesis-relat
39	28	96.6	177	2	S04728	pathogenesis-relat
40	28	96.6	177	2	T01705	hypothetical prote
41	28	96.6	179	2	S22531	pathogenesis-relat
42	28	96.6	185	2	C83644	conserved hypothet
43	28	96.6	187	2	A82746	conserved hypothet
44	28	96.6	190	2	AG0030	conserved hypothet
45	28	96.6	191	2	AH3005	conserved hypothet

ALIGNMENTS

RESULT 1
C39925
hypothetical protein 2 - equine arteritis virus
C;Species: equine arteritis virus
C;Date: 14-Feb-1992 #sequence_revision 14-Feb-1992 #text_change 09-Jul-2004
C;Accession: C39925
R;Den Boon, J.A.; Snijder, E.J.; Chirnside, E.D.; De Vries, A.A.F.; Horzinek, M.C.; Spaar
J. Virol. 65, 2910-2920, 1991
A;Title: Equine arteritis virus is not a togavirus but belongs to the coronaviruslike sup
A;Reference number: A39925; MUID:91237805; PMID:1851863
A;Accession: C39925
A;Status: preliminary
A;Molecule type: genomic RNA
A;Residues: 1-227 <DEN>
A;Cross-references: UNIPROT:P28992; UNIPARC:UPI000011F47A; EMBL:X53459; NID:g62065; PIDN
C;Superfamily: equine arteritis virus hypothetical protein 2

Query Match 100.0%; Score 29; DB 2; Length 227;
Best Local Similarity 80.0%; Pred. No. 82;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 GPXTW 8
|| ||
Db 158 GPATW 162

RESULT 2
G87286
conserved hypothetical protein CC0304 [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C;Accession: G87286
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.E
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolona
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: G87286
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-237 <STO>
A;Cross-references: UNIPROT:Q9ABC6; UNIPARC:UPI00000C6FD0; GB:AE005673; NID:g13421447; P
C;Genetics:
A;Gene: CC0304

Query Match 100.0%; Score 29; DB 2; Length 237;
Best Local Similarity 80.0%; Pred. No. 86;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 GPXTW 8
|| ||

```
Db         64 GPATW 68


RESULT 3
B86784
hypothetical protein ynaE [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C;Species: Lactococcus lactis subsp. lactis
C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004
C;Accession: B86784
R;Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarme, K.; Weissenbach, J.; Ehrli
Genome Res. 11, 731-753, 2001
A;Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ss
A;Reference number: A86625; MUID:21235186; PMID:11337471
A;Accession: B86784
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-246 <STO>
A;Cross-references: UNIPROT:Q9CG36; UNIPARC:UPI00000C69F1; GB:AE005176; PID:g12724250; P
A;Experimental source: strain IL1403
C;Genetics:
A;Gene: ynaE

  Query Match             100.0%;  Score 29;  DB 2;  Length 246;
  Best Local Similarity   80.0%;  Pred. No. 89;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 GPXTW 8
              || ||
Db        115 GPSTW 119


RESULT 4
T48742
hypothetical protein 8D4.160 [imported] - Neurospora crassa
C;Species: Neurospora crassa
C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 19-May-2000
C;Accession: T48742
R;Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, April 2000
A;Reference number: Z24541
A;Accession: T48742
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-263 <SCH>
A;Cross-references: UNIPARC:UPI0000179476; EMBL:AL353819; GSPDB:GN00112; NCSP:8D4.160
A;Experimental source: cosmid contig 8D4; strain 74
C;Genetics:
A;Gene: NCSP:8D4.160
A;Map position: 2
A;Introns: 32/3; 76/1; 133/1
C;Superfamily: Neurospora crassa hypothetical protein 8D4.160

  Query Match             100.0%;  Score 29;  DB 2;  Length 263;
  Best Local Similarity   80.0%;  Pred. No. 95;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 GPXTW 8
              || ||
Db        252 GPSTW 256


RESULT 5
AC3647
cellobiose phosphotransferase system celC [imported] - Brucella melitensis (strain 16M)
C;Species: Brucella melitensis
C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C;Accession: AC3647
R;DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
.; Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitens
A;Reference number: AD3252; PMID:11756688
A;Accession: AC3647
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-279 <KUR>
A;Cross-references: UNIPROT:Q8YB01; UNIPARC:UPI0000058739; GB:AE008918; PIDN:AAL54342.1;
A;Experimental source: strain 16M
C;Genetics:
A;Gene: BMEII1100
A;Map position: II

  Query Match             100.0%;  Score 29;  DB 2;  Length 279;
  Best Local Similarity   80.0%;  Pred. No. 1e+02;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 GPXTW 8
              || ||
Db        164 GPATW 168


RESULT 6
B87356
sugar ABC transporter, permease protein CC0861 [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C;Accession: B87356
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
 B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolona
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: B87356
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-332 <STO>
A;Cross-references: UNIPROT:Q9A9V0; UNIPARC:UPI00000C71C6; GB:AE005673; NID:g13422120; PI
C;Genetics:
A;Gene: CC0861
C;Superfamily: l-arabinose transport system permease araH

  Query Match             100.0%;  Score 29;  DB 2;  Length 332;
  Best Local Similarity   80.0%;  Pred. No. 1.2e+02;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 GPXTW 8
              || ||
Db        104 GPATW 108


RESULT 7
JC6082
proximal sequence element-binding transcription factor delta chain - human
C;Species: Homo sapiens (man)
C;Date: 13-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C;Accession: JC6082; PC6030; G02375
R;Yoon, J.B.; Roeder, R.G.
Mol. Cell. Biol. 16, 1-9, 1996
A;Title: Cloning of two proximal sequence element-binding transcription factor subunits
and interact with the TATA-binding protein.
A;Reference number: JC6081; MUID:96104548; PMID:8524284
A;Accession: JC6082
A;Molecule type: mRNA
A;Residues: 1-334 <YOO1>
A;Cross-references: UNIPROT:Q13487; UNIPARC:UPI000016A2B2; GB:U44755; NID:g1174204; PIDN
A;Accession: PC6030
A;Molecule type: protein
A;Residues: 81-102;151-175;307-321 <YOO2>
A;Cross-references: UNIPARC:UPI0000179AE6; UNIPARC:UPI0000179AE7; UNIPARC:UPI0000179AE8
R;Henry, R.W.
submitted to the EMBL Data Library, January 1996
A;Reference number: H01137
A;Accession: G02375
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
```

A;Residues: 1-117,'L',119-334 <HEN>
A;Cross-references: UNIPARC:UPI0000000CBC; EMBL:U44898; NID:g1174257; PIDN:AAB06230.1; P
C;Comment: This factor is highly acidic. It recognizes the proximal sequence elements, p
basal transcription, communicates with both class II and class III general transcription
C;Genetics:
A;Gene: SNAP45
C;Keywords: transcription factor
F;141-196/Region: proline-rich

Query Match 100.0%; Score 29; DB 2; Length 334;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 GPXTW 8
|| ||
Db 20 GPATW 24

RESULT 8
T42623
probable sugar transport protein - fission yeast (Schizosaccharomyces pombe) (fragment)
C;Species: Schizosaccharomyces pombe
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 21-Jul-2000
C;Accession: T42623
R;Yoshioka, S.; Kato, K.; Nakai, K.; Okayama, H.; Nojima, H.
DNA Res. 4, 363-369, 1997
A;Title: Identification of open reading frames in Schizosaccharomyces pombe cDNAs.
A;Reference number: Z17323; MUID:98162722; PMID:9501991
A;Accession: T42623
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-371 <YOS>
A;Cross-references: UNIPARC:UPI00001690FF; EMBL:D89179; NID:g1749565; PIDN:BAA13841.1; P
A;Experimental source: strain PR745
C;Superfamily: maltose transport protein MAL61
C;Keywords: sugar transport; transmembrane protein

Query Match 100.0%; Score 29; DB 2; Length 371;
Best Local Similarity 80.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 GPXTW 8
|| ||
Db 75 GPATW 79

RESULT 9
T44587
cytochrome P450 homolog [imported] - Streptomyces fradiae
C;Species: Streptomyces fradiae
C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004
C;Accession: T44587
R;Bate, N.; Butler, A.R.; Gandecha, A.R.; Cundliffe, E.
Chem. Biol. 6, 617-624, 1999
A;Title: Multiple regulatory genes in the tylosin-biosynthetic cluster of Streptomyces f
A;Reference number: Z22801; MUID:99398833; PMID:10467127
A;Accession: T44587
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-433 <BAT>
A;Cross-references: UNIPROT:Q9XCC6; UNIPARC:UPI00000AF273; EMBL:AF145049; PIDN:AAD40802.
A;Experimental source: strain T59235
C;Superfamily: Bacillus cytochrome P450 CYP106; cytochrome P450 homology
F;262-399/Domain: cytochrome P450 homology <P45>

Query Match 100.0%; Score 29; DB 2; Length 433;
Best Local Similarity 80.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 GPXTW 8
|| ||
Db 52 GPATW 56

RESULT 10
S63383
POP2 protein - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein N3470; protein YNR052c
C;Species: Saccharomyces cerevisiae
C;Date: 27-Apr-1996 #sequence_revision 03-May-1996 #text_change 09-Jul-2004
C;Accession: S63383; S35997; S35996; S36929; S27438
R;Pohl, T.M.
submitted to the Protein Sequence Database, April 1996
A;Reference number: S63346
A;Accession: S63383
A;Molecule type: DNA
A;Residues: 1-433 <POH>
A;Cross-references: UNIPROT:P39008; UNIPARC:UPI0000052DFA; EMBL:Z71667; NID:g1302567; PID
A;Experimental source: strain S288C
R;Sakai, A.; Chibazakura, T.; Shimizu, Y.; Hishinuma, F.
Nucleic Acids Res. 20, 6227-6233, 1992
A;Title: Molecular analysis of POP2 gene, a gene required for glucose-derepression of ge
A;Reference number: S35996; MUID:93117094; PMID:1475183
A;Accession: S35997
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-80,82-411,'M',413-433 <SAK>
A;Cross-references: UNIPARC:UPI000017B2FC; GB:D12807
A;Experimental source: strain S288C
A;Accession: S35996
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-40,'Q',42-91,'QQQQQQQQQQQQQQQQ',92-111,117-277,'S',279-433 <SAW>
A;Cross-references: UNIPARC:UPI000017B2FD; GB:D12808
A;Experimental source: strain A364A
R;Sakai, A.; Chibazakura, T.; Shimizu, Y.; Hishinuma, F.
submitted to the EMBL Data Library, August 1992
A;Reference number: S36929
A;Accession: S36929
A;Molecule type: DNA
A;Residues: 1-91,'QQQQQQQQQQQQQQQQ',92-111,117-277,'S',279-433 <SA2>
A;Cross-references: UNIPARC:UPI0000168D95; GB:D12808; NID:g218462; PID:d1002742; PID:g21
A;Experimental source: strain A364A
R;Cusick, M.E.
submitted to the EMBL Data Library, March 1992
A;Reference number: S27437
A;Accession: S27438
A;Molecule type: DNA
A;Residues: 213-433 <CUS>
A;Cross-references: UNIPARC:UPI0000168D3A; EMBL:M88607; NID:g172079; PID:g172080
C;Genetics:
A;Gene: SGD:POP2; CAF1
A;Cross-references: SGD:S0005335; MIPS:YNR052c
A;Map position: 14R

Query Match 100.0%; Score 29; DB 2; Length 433;
Best Local Similarity 80.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 GPXTW 8
|| ||
Db 240 GPSTW 244

RESULT 11
S36507
L2 protein - human papillomavirus type 30
C;Species: human papillomavirus type 30
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C;Accession: S36507
R;Delius, H.; Hofmann, B.
submitted to the EMBL Data Library, August 1993
A;Description: Primer-directed sequencing of human papillomavirus types.
A;Reference number: S36469
A;Accession: S36507
A;Molecule type: DNA

A;Residues: 1-463 <DEL>
A;Cross-references: UNIPROT:P36756; UNIPARC:UPI0000138901; EMBL:X74474; NID:g396973; PID
C;Superfamily: papillomavirus L2 protein
C;Keywords: late protein

Query Match 100.0%; Score 29; DB 2; Length 463;
Best Local Similarity 80.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 GPXTW 8
|| ||
Db 411 GPTTW 415

RESULT 12
S36582
L2 protein - human papillomavirus type 56
C;Species: human papillomavirus type 56
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C;Accession: S36582
R;Delius, H.; Hofmann, B.
submitted to the EMBL Data Library, August 1993
A;Description: Primer-directed sequencing of human papillomavirus types.
A;Reference number: S36469
A;Accession: S36582
A;Molecule type: DNA
A;Residues: 1-464 <DEL>
A;Cross-references: UNIPROT:P36765; UNIPARC:UPI0000138919; EMBL:X74483; NID:g397053; PID
C;Superfamily: papillomavirus L2 protein
C;Keywords: late protein

Query Match 100.0%; Score 29; DB 2; Length 464;
Best Local Similarity 80.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 GPXTW 8
|| ||
Db 412 GPSTW 416

RESULT 13
H84137
hypothetical protein BH3904 [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C;Accession: H84137
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: H84137
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-475 <STO>
A;Cross-references: UNIPROT:Q9K628; UNIPARC:UPI00000C432F; GB:AP001520; GB:BA000004; NID
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH3904

Query Match 100.0%; Score 29; DB 2; Length 475;
Best Local Similarity 80.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 GPXTW 8
|| ||
Db 159 GPSTW 163

RESULT 14
D70703
tldD homolog Rv2315c - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 29-Jan-1999 #sequence_revision 29-Jan-1999 #text_change 09-Jul-2004
C;Accession: D70703
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome s
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: D70703
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-505 <COL>
A;Cross-references: UNIPROT:P71897; UNIPARC:UPI00000D11F8; GB:Z79702; GB:AL123456; NID:g3
A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: Rv2315c
C;Superfamily: Escherichia coli tldD protein

Query Match 100.0%; Score 29; DB 1; Length 505;
Best Local Similarity 80.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 GPXTW 8
|| ||
Db 459 GPSTW 463

RESULT 15
D90207
conserved hypothetical protein [imported] - Sulfolobus solfataricus
C;Species: Sulfolobus solfataricus
C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004
C;Accession: D90207
R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-W
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A;Description: Sulfolobus solfataricus complete genome.
A;Reference number: A99139
A;Accession: D90207
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-506 <KUR>
A;Cross-references: UNIPROT:Q9UWZ8; UNIPARC:UPI0000064A6F; GB:AE006641; NID:g13813769; PI
C;Genetics:
A;Gene: SSO0604

Query Match 100.0%; Score 29; DB 2; Length 506;
Best Local Similarity 80.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 GPXTW 8
|| ||
Db 153 GPSTW 157

Search completed: March 31, 2006, 16:37:20
Job time : 7.21891 secs

OM protein - protein search, using sw model

Run on: March 31, 2006, 16:09:36 ; Search time 37.4627 Seconds
(without alignments)
188.328 Million cell updates/sec

Title: US-10-609-217-124
Perfect score: 29
Sequence: 1 XXXGPXTWXX 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	29	100.0	69	2	Q82I26_STRAW	Q82i26 streptomyce
2	29	100.0	74	2	Q50168_MYCLE	Q50168 mycobacteri
3	29	100.0	78	2	Q7YWR6_CAEEL	Q7ywr6 caenorhabdi
4	29	100.0	89	2	Q6N846_RHOPA	Q6n846 rhodopseudo
5	29	100.0	96	2	Q654Q1_ORYSA	Q654q1 oryza sativ
6	29	100.0	97	2	Q764H5_TYLCV	Q764h5 tomato yell
7	29	100.0	105	2	Q4HYB3_GIBZE	Q4hyb3 gibberella
8	29	100.0	111	2	Q4R9N6_TETNG	Q4r9n6 tetraodon n
9	29	100.0	135	2	Q949Z8_ARATH	Q949z8 arabidopsis
10	29	100.0	136	2	Q5SI79_THET8	Q5si79 thermus the
11	29	100.0	136	2	Q72IK1_THET2	Q72ik1 thermus the
12	29	100.0	139	2	Q8KGS9_RHILO	Q8kgs9 rhizobium l
13	29	100.0	147	2	Q8TVJ1_METKA	Q8tvj1 methanopyru
14	29	100.0	152	2	O15113_HUMAN	O15113 homo sapien
15	29	100.0	155	2	Q6Z8V0_ORYSA	Q6z8v0 oryza sativ
16	29	100.0	156	2	Q629Y4_BURMA	Q629y4 burkholderi
17	29	100.0	164	2	Q4H1M0_9BACT	Q4h1m0 uncultured
18	29	100.0	165	2	Q4H1M8_9BACT	Q4h1m8 uncultured
19	29	100.0	179	2	Q6AHM5_LEIXX	Q6ahm5 leifsonia x
20	29	100.0	181	2	Q6P2N5_HUMAN	Q6p2n5 homo sapien
21	29	100.0	186	2	Q603G8_METCA	Q603g8 methylococc
22	29	100.0	196	2	Q8U5X1_AGRT5	Q8u5x1 agrobacteri
23	29	100.0	199	2	Q9LK79_ARATH	Q9lk79 arabidopsis
24	29	100.0	205	2	Q8BEF2_9PICO	Q8bef2 foot-and-mo
25	29	100.0	207	2	Q6T316_HUMAN	Q6t316 homo sapien
26	29	100.0	208	2	Q8BED4_9PICO	Q8bed4 foot-and-mo
27	29	100.0	208	2	Q8BED9_9PICO	Q8bed9 foot-and-mo
28	29	100.0	209	2	Q98A43_RHILO	Q98a43 rhizobium l
29	29	100.0	209	2	Q8QTX4_9PICO	Q8qtx4 foot-and-mo
30	29	100.0	209	2	Q8QTX6_9PICO	Q8qtx6 foot-and-mo
31	29	100.0	209	2	Q8QTX7_9PICO	Q8qtx7 foot-and-mo
32	29	100.0	209	2	Q8QTX8_9PICO	Q8qtx8 foot-and-mo
33	29	100.0	209	2	Q8BEB4_9PICO	Q8beb4 foot-and-mo
34	29	100.0	209	2	Q8BEB5_9PICO	Q8beb5 foot-and-mo
35	29	100.0	209	2	Q8BEC4_9PICO	Q8bec4 foot-and-mo
36	29	100.0	210	2	Q4WZ27_ASPFU	Q4wz27 aspergillus
37	29	100.0	210	2	Q7Q2J1_ANOGA	Q7q2j1 anopheles g
38	29	100.0	210	2	Q8QTX1_9PICO	Q8qtx1 foot-and-mo
39	29	100.0	210	2	Q8QTX3_9PICO	Q8qtx3 foot-and-mo
40	29	100.0	210	2	Q8QTY7_9PICO	Q8qty7 foot-and-mo
41	29	100.0	210	2	Q8BEB3_9PICO	Q8beb3 foot-and-mo
42	29	100.0	210	2	Q8BED8_9PICO	Q8bed8 foot-and-mo
43	29	100.0	211	2	Q8QTW9_9PICO	Q8qtw9 foot-and-mo
44	29	100.0	211	2	Q8QTY3_9PICO	Q8qty3 foot-and-mo
45	29	100.0	211	2	Q8QTY5_9PICO	Q8qty5 foot-and-mo

ALIGNMENTS

RESULT 1
Q82I26_STRAW
ID Q82I26_STRAW PRELIMINARY; PRT; 69 AA.
AC Q82I26;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=SAV3332;
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=22608306; PubMed=12692562; DOI=10.1038/nbt820;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis.";
RL Nat. Biotechnol. 21:526-531(2003).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
DR EMBL; BA000030; BAC71043.1; -; Genomic_DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 69 AA; 7190 MW; C66534F3269F2379 CRC64;

Query Match 100.0%; Score 29; DB 2; Length 69;
Best Local Similarity 80.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          4 GPXTW 8
              || ||
Db         58 GPTTW 62
```

RESULT 2
Q50168_MYCLE
ID Q50168_MYCLE PRELIMINARY; PRT; 74 AA.
AC Q50168;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE U296u.

```
OS   Mycobacterium leprae.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX   NCBI_TaxID=1769;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Smith D.R.;
RL   Submitted (APR-1995) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RA   Robison K.;
RL   Submitted (SEP-1994) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U15187; AAA63114.1; -; Genomic_DNA.
SQ   SEQUENCE   74 AA;  8949 MW;  687DABA58947E513 CRC64;

  Query Match             100.0%;  Score 29;  DB 2;  Length 74;
  Best Local Similarity   80.0%;  Pred. No. 2e+02;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 GPXTW 8
              || ||
Db         21 GPTTW 25


RESULT 3
Q7YWR6_CAEEL
ID   Q7YWR6_CAEEL PRELIMINARY;      PRT;    78 AA.
AC   Q7YWR6;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Hypothetical protein T24F1.7.
GN   ORFNames=T24F1.7;
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC   STRAIN=Bristol N2;
RX   MEDLINE=99069613; PubMed=9851916;
RG   The C. elegans sequencing consortium;
RT   "Genome sequence of the nematode C. elegans: a platform for
RT   investigating biology.";
RL   Science 282:2012-2018(1998).
DR   EMBL; Z49912; CAE17988.1; -; Genomic_DNA.
DR   Ensembl; T24F1.7; Caenorhabditis elegans.
DR   WormBase; WBGene00011999; T24F1.7.
DR   WormPep; T24F1.7; CE35029.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   78 AA;  8832 MW;  59377E53D82DCC3A CRC64;

  Query Match             100.0%;  Score 29;  DB 2;  Length 78;
  Best Local Similarity   80.0%;  Pred. No. 2.1e+02;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 GPXTW 8
              || ||
Db         13 GPTTW 17


RESULT 4
Q6N846_RHOPA
ID   Q6N846_RHOPA PRELIMINARY;      PRT;    89 AA.
AC   Q6N846;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Hypothetical protein.
GN   OrderedLocusNames=RPA2058;
OS   Rhodopseudomonas palustris.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC   Bradyrhizobiaceae; Rhodopseudomonas.
OX   NCBI_TaxID=1076;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=CGA009 / ATCC BAA-98;
RX   PubMed=14704707; DOI=10.1038/nbt923;
RA   Larimer F.W., Chain P., Hauser L., Lamerdin J.E., Malfatti S., Do L.,
RA   Land M.L., Pelletier D.A., Beatty J.T., Lang A.S., Tabita F.R.,
RA   Gibson J.L., Hanson T.E., Bobst C., Torres y Torres J.L., Peres C.,
RA   Harrison F.H., Gibson J., Harwood C.S.;
RT   "Complete genome sequence of the metabolically versatile
RT   photosynthetic bacterium Rhodopseudomonas palustris.";
RL   Nat. Biotechnol. 22:55-61(2004).
DR   EMBL; BX572599; CAE27499.1; -; Genomic_DNA.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   89 AA;  10363 MW;  D98873C13B498888 CRC64;

  Query Match             100.0%;  Score 29;  DB 2;  Length 89;
  Best Local Similarity   80.0%;  Pred. No. 2.4e+02;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 GPXTW 8
              || ||
Db         67 GPATW 71


RESULT 5
Q654Q1_ORYSA
ID   Q654Q1_ORYSA PRELIMINARY;      PRT;    96 AA.
AC   Q654Q1;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Hypothetical protein OSJNBa0091G06.13.
GN   Name=OSJNBa0091G06.13;
OS   Oryza sativa (japonica cultivar-group).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=39947;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Sasaki T., Matsumoto T., Yamamoto K.;
RT   "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 6, BAC
RT   clone:OSJNBa0091G06.";
RL   Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AP004651; BAD45716.1; -; Genomic_DNA.
DR   Gramene; Q654Q1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   96 AA;  10350 MW;  3382221FCDCE8539 CRC64;

  Query Match             100.0%;  Score 29;  DB 2;  Length 96;
  Best Local Similarity   80.0%;  Pred. No. 2.6e+02;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 GPXTW 8
              || ||
Db         61 GPSTW 65


RESULT 6
Q764H5_TYLCV
ID   Q764H5_TYLCV PRELIMINARY;      PRT;    97 AA.
AC   Q764H5;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   C4.
OS   Tomato yellow leaf curl virus (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10832;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Miy;
```

```
RA   Ueda S., Kimura T., Onuki M., Hanada K., Iwanami T.;
RT   "Three distinct groups of isolates of Tomato yellow leaf curl virus in
RT   Japan and construction of an infectious clone.";
RL   J. Gen. Plant Pathol. 70:232-238(2004).
DR   EMBL; AB116629; BAD07419.1; -; Genomic_DNA.
DR   InterPro; IPR002488; Gemini_C4.
DR   Pfam; PF01492; Gemini_C4; 1.
SQ   SEQUENCE   97 AA;  11099 MW;  0ADE53DBBCED04B4 CRC64;

  Query Match             100.0%;  Score 29;  DB 2;  Length 97;
  Best Local Similarity   80.0%;  Pred. No. 2.6e+02;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 GPXTW 8
              || ||
Db         22 GPSTW 26


RESULT 7
Q4HYB3_GIBZE
ID   Q4HYB3_GIBZE PRELIMINARY;      PRT;   105 AA.
AC   Q4HYB3;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Predicted protein.
GN   ORFNames=FG10045.1;
OS   Gibberella zeae PH-1.
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC   Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
OX   NCBI_TaxID=229533;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=PH-1;
RA   Birren B., Nusbaum C., Abouelleil A., Allen N., Anderson S.,
RA   Arachchi H.M., Barna N., Bastien V., Bloom T., Boguslavkiy L.,
RA   Boukhgalter B., Butler J., Calvo S.E., Camarata J., Chang J.,
RA   Choepel Y., Collymore A., Cook A., Cooke P., Corum B., DeArellano K.,
RA   Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
RA   Erickson J., Faro S., Ferreira P., FitzGerald M., Gage D., Galagan J.,
RA   Gardyna S., Gnerre S., Graham L., Grand-Pierre N., Hafez N.,
RA   Hagopian D., Hagos B., Hall J., Horton L., Hulme W., Iliev I.,
RA   Jaffe D., Johnson R., Jones C., Kamal M., Kamat A., Karatas A.,
RA   Kells C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
RA   Ma L.-J., Mabbitt R., MacLean C., Macdonald P., Major J., Manning J.,
RA   Matthews C., Mauceli E., McCarthy M., Meldrim J., Meneus L.,
RA   Mihova T., Mlenga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
RA   Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neil D.,
RA   Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,
RA   Rachupka A., Ramasamy U., Raymond C., Retta R., Rise C., Rogov P.,
RA   Roman J., Schauer S., Schupback R., Seaman S., Severy P., Smirnov S.,
RA   Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,
RA   Talamas J., Tesfaye S., Theodore J., Topham K., Travers M.,
RA   Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,
RA   Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
RA   Lander E.;
RT   "Fusarium graminearum genome sequence.";
RL   Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; AACM01000417; EAA70361.1; -; Genomic_DNA.
SQ   SEQUENCE   105 AA;  11626 MW;  7B4ADB5173C592A2 CRC64;

  Query Match             100.0%;  Score 29;  DB 2;  Length 105;
  Best Local Similarity   80.0%;  Pred. No. 2.8e+02;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 GPXTW 8
              || ||
Db         21 GPTTW 25


RESULT 8
Q4R9N6_TETNG
ID   Q4R9N6_TETNG PRELIMINARY;      PRT;   111 AA.
AC   Q4R9N6;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Chromosome undetermined SCAF25737, whole genome shotgun sequence.
DE   (Fragment).
GN   ORFNames=GSTENG00037755001;
OS   Tetraodon nigroviridis (Green puffer).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC   Tetradontoidea; Tetraodontidae; Tetraodon.
OX   NCBI_TaxID=99883;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA   Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA   Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA   Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA   Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA   Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA   Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA   Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA   Kellis M., Volff JN., Guigo R., Zody M.C., Mesirov J.,
RA   Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA   Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA   Wincker P., Lander E.S., Weissenbach J., Roest Crollius H.;
RT   "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT   the early vertebrate proto-karyotype.";
RL   Nature 431:946-957(2004).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RG   Genoscope; Whitehead Institute Centre for Genome Research;
RL   Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; CAAE01025737; CAG14897.1; -; Genomic_DNA.
FT   NON_TER       1       1
SQ   SEQUENCE   111 AA;  13292 MW;  A7231EDF0A18F377 CRC64;

  Query Match             100.0%;  Score 29;  DB 2;  Length 111;
  Best Local Similarity   80.0%;  Pred. No. 3e+02;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 GPXTW 8
              || ||
Db        100 GPSTW 104


RESULT 9
Q949Z8_ARATH
ID   Q949Z8_ARATH PRELIMINARY;      PRT;   135 AA.
AC   Q949Z8;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE   Hypothetical protein At1g36980.
GN   Name=At1g36980;
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
RA   Goldsmith A.D., Lee J.M., Quach H.L., Toriumi M., Yu G., Bowser L.,
RA   Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA   Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
```

```
RA   Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,
RA   Seki M., Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
RA   Theologis A.;
RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RA   Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Etgu P., Lee J.M.,
RA   Toriumi M., Yu G., Brooks S., Chao Q., Chen H., Karlin-Neumann G.,
RA   Kim C., Lam B., Miranda M., Nguyen M., Palm C.J., Shinn P.,
RA   Southwick A., Davis R.W., Ecker J.R., Theologis A.;
RL   Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY050785; AAK92720.1; -; mRNA.
DR   EMBL; AY079344; AAL85075.1; -; mRNA.
DR   InterPro; IPR007919; UPF0220.
DR   Pfam; PF05255; UPF0220; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   135 AA;  14959 MW;  1238AE0D5FB58DDB CRC64;

  Query Match             100.0%;  Score 29;  DB 2;  Length 135;
  Best Local Similarity   80.0%;  Pred. No. 3.6e+02;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 GPXTW 8
              || ||
Db        106 GPSTW 110


RESULT 10
Q5SI79_THET8
ID   Q5SI79_THET8 PRELIMINARY;      PRT;   136 AA.
AC   Q5SI79;
DT   01-FEB-2005 (TrEMBLrel. 29, Created)
DT   01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT   01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE   Hypothetical protein TTHA1495.
GN   OrderedLocusNames=TTHA1495;
OS   Thermus thermophilus (strain HB8 / ATCC 27634 / DSM 579).
OC   Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;
OC   Thermus.
OX   NCBI_TaxID=300852;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=HB8;
RA   Masui R., Kurokawa K., Nakagawa N., Tokunaga F., Koyama Y.,
RA   Shibata T., Oshima T., Yokoyama S., Yasunaga T., Kuramitsu S.;
RT   "Complete genome sequence of Thermus thermophilus HB8.";
RL   Submitted (NOV-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AP008226; BAD71318.1; -; Genomic_DNA.
DR   InterPro; IPR010432; RDD.
DR   Pfam; PF06271; RDD; 1.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   136 AA;  15192 MW;  73FF945DB6C25726 CRC64;

  Query Match             100.0%;  Score 29;  DB 2;  Length 136;
  Best Local Similarity   80.0%;  Pred. No. 3.6e+02;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 GPXTW 8
              || ||
Db         37 GPTTW 41


RESULT 11
Q72IK1_THET2
ID   Q72IK1_THET2 PRELIMINARY;      PRT;   136 AA.
AC   Q72IK1;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Hypothetical membrane spanning protein.
GN   OrderedLocusNames=TTC1131;
OS   Thermus thermophilus (strain HB27 / ATCC BAA-163 / DSM 7039).
OC   Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;
OC   Thermus.
OX   NCBI_TaxID=262724;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   PubMed=15064768; DOI=10.1038/nbt956;
RA   Henne A., Brueggemann H., Raasch C., Wiezer A., Hartsch T.,
RA   Liesegang H., Johann A., Lienard T., Gohl O., Martinez-Arias R.,
RA   Jacobi C., Starkuviene V., Schlenczeck S., Dencker S., Huber R.,
RA   Klenk H.-P., Kramer W., Merkl R., Gottschalk G., Fritz H.-J.;
RT   "The genome sequence of the extreme thermophile Thermus
RT   thermophilus.";
RL   Nat. Biotechnol. 22:547-553(2004).
DR   EMBL; AE017305; AAS81473.1; -; Genomic_DNA.
DR   InterPro; IPR010432; RDD.
DR   Pfam; PF06271; RDD; 1.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   136 AA;  15192 MW;  73FF945DB6C25726 CRC64;

  Query Match             100.0%;  Score 29;  DB 2;  Length 136;
  Best Local Similarity   80.0%;  Pred. No. 3.6e+02;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 GPXTW 8
              || ||
Db         37 GPTTW 41


RESULT 12
Q8KGS9_RHILO
ID   Q8KGS9_RHILO PRELIMINARY;      PRT;   139 AA.
AC   Q8KGS9;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Hypothetical protein msi151.
GN   Name=msi151;
OS   Rhizobium loti (Mesorhizobium loti).
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC   Phyllobacteriaceae; Mesorhizobium.
OX   NCBI_TaxID=381;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=R7A;
RX   MEDLINE=21999272; PubMed=12003951;
RX   DOI=10.1128/JB.184.11.3086-3095.2002;
RA   Sullivan J.T., Trzebiatowski J.R., Cruickshank R.W., Gouzy J.,
RA   Brown S.D., Elliot R.M., Fleetwood D.J., McCallum N.G., Rossbach U.,
RA   Stuart G.S., Weaver J.E., Webby R.J., de Bruijn F.J., Ronson C.W.;
RT   "Comparative sequence analysis of the symbiosis island of
RT   Mesorhizobium loti strain R7A.";
RL   J. Bacteriol. 184:3086-3095(2002).
DR   EMBL; AL672113; CAD31556.1; -; Genomic_DNA.
KW   Hypothetical protein.
SQ   SEQUENCE   139 AA;  14765 MW;  5587C028235AB423 CRC64;

  Query Match             100.0%;  Score 29;  DB 2;  Length 139;
  Best Local Similarity   80.0%;  Pred. No. 3.7e+02;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 GPXTW 8
              || ||
Db         35 GPATW 39


RESULT 13
Q8TVJ1_METKA
ID   Q8TVJ1_METKA PRELIMINARY;      PRT;   147 AA.
AC   Q8TVJ1;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Transcriptional regulator of the riboflavin/FAD biosynthetic
DE   operon.
```

```
GN   OrderedLocusNames=MK1398;
OS   Methanopyrus kandleri.
OC   Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;
OC   Methanopyrus.
OX   NCBI_TaxID=2320;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC   STRAIN=AV19 / DSM 6324 / JCM 9639;
RX   MEDLINE=21927647; PubMed=11930014; DOI=10.1073/pnas.032671499;
RA   Slesarev A.I., Mezhevaya K.V., Makarova K.S., Polushin N.N.,
RA   Shcherbinina O.V., Shakhova V.V., Belova G.I., Aravind L.,
RA   Natale D.A., Rogozin I.B., Tatusov R.L., Wolf Y.I., Stetter K.O.,
RA   Malykh A.G., Koonin E.V., Kozyavkin S.A.;
RT   "The complete genome of hyperthermophile Methanopyrus kandleri AV19
RT   and monophyly of archaeal methanogens.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).
DR   EMBL; AE010432; AAM02611.1; -; Genomic_DNA.
DR   InterPro; IPR002834; DUF120.
DR   Pfam; PF01982; DUF120; 1.
DR   ProDom; PD015839; DUF120; 1.
KW   Complete proteome.
SQ   SEQUENCE   147 AA;  16880 MW;  1FBFA65B4D2AB012 CRC64;

  Query Match             100.0%;  Score 29;  DB 2;  Length 147;
  Best Local Similarity   80.0%;  Pred. No. 3.9e+02;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 GPXTW 8
              || ||
Db        129 GPSTW 133


RESULT 14
O15113_HUMAN
ID   O15113_HUMAN PRELIMINARY;      PRT;   152 AA.
AC   O15113;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Hypothetical protein (Fragment).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Liver;
RA   Feng S.J., McKeehan W.L.;
RL   Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF015910; AAB64298.1; -; mRNA.
KW   Hypothetical protein.
FT   NON_TER       1      1
FT   NON_TER     152    152
SQ   SEQUENCE   152 AA;  16322 MW;  602423BBE49720F4 CRC64;

  Query Match             100.0%;  Score 29;  DB 2;  Length 152;
  Best Local Similarity   80.0%;  Pred. No. 4e+02;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 GPXTW 8
              || ||
Db          9 GPTTW 13


RESULT 15
Q6Z8V0_ORYSA
ID   Q6Z8V0_ORYSA PRELIMINARY;      PRT;   155 AA.
AC   Q6Z8V0;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE   Hypothetical protein P0686H11.7 (Hypothetical protein
DE   P0605H02.47).
GN   Name=P0686H11.7; Synonyms=P0605H02.47;
OS   Oryza sativa (japonica cultivar-group).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=39947;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Sasaki T., Matsumoto T., Yamamoto K.;
RL   Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RA   Sasaki T., Matsumoto T., Yamamoto K.;
RT   "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 8, PAC
RT   clone:P0605H02.";
RL   Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AP004762; BAD10001.1; -; Genomic_DNA.
DR   EMBL; AP004620; BAD09679.1; -; Genomic_DNA.
DR   Gramene; Q6Z8V0; -.
KW   Hypothetical protein.
SQ   SEQUENCE   155 AA;  17235 MW;  0E5A9876140B7261 CRC64;

  Query Match             100.0%;  Score 29;  DB 2;  Length 155;
  Best Local Similarity   80.0%;  Pred. No. 4.1e+02;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 GPXTW 8
              || ||
Db        151 GPSTW 155


Search completed: March 31, 2006, 16:35:14
Job time : 39.4627 secs
```

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: March 31, 2006, 16:09:06 ; Search time 188.93 Seconds
(without alignments)
113.955 Million cell updates/sec

Title: US-10-609-217-339
Perfect score: 306
Sequence: 1 GGGGGGGTYSCHFGPLTWVC.....GGTYSCHFGPLTWVCKPQGG 49

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	306	100.0	49	5	ABB73392	Abb73392 EPO-mimet
2	306	100.0	50	3	AAB17283	Aab17283 EPO-mimet
3	306	100.0	277	3	AAB16967	Aab16967 Fc-EMP-EM
4	306	100.0	277	5	ABB73418	Abb73418 Fc-EMP-EM
5	276	90.2	49	5	ABB73393	Abb73393 EPO-mimet
6	276	90.2	50	3	AAB17284	Aab17284 EPO-mimet
7	276	90.2	57	3	AAB17314	Aab17314 EMP-EMP-F
8	276	90.2	57	5	ABB73408	Abb73408 EMP-EMP g
9	276	90.2	277	3	AAB16966	Aab16966 EMP-EMP-F
10	276	90.2	278	5	ABB73417	Abb73417 EMP-EMP-F
11	259	84.6	70	7	ADJ72562	Adj72562 EPO mimet
12	249	81.4	47	3	AAB17040	Aab17040 EPO-mimet
13	249	81.4	47	8	ADJ52198	Adj52198 CH1 delet
14	249	81.4	47	8	ADJ51160	Adj51160 CH1 delet
15	240	78.4	40	3	AAB17036	Aab17036 EPO-mimet
16	240	78.4	40	5	ABB72819	Abb72819 Erythropo
17	240	78.4	40	8	ADJ52195	Adj52195 CH1 delet
18	239.5	78.3	41	3	AAB17037	Aab17037 EPO-mimet
19	239.5	78.3	41	5	ABB72820	Abb72820 Erythropo
20	239.5	78.3	41	7	ADJ72559	Adj72559 EPO mimet
21	239.5	78.3	41	8	ADJ51157	Adj51157 CH1 delet
22	239.5	78.3	46	3	AAB17039	Aab17039 EPO-mimet
23	239.5	78.3	46	5	ABB72822	Abb72822 Erythropo
24	235	76.8	47	5	ABB72823	Abb72823 Erythropo
25	231	75.5	39	3	AAB17312	Aab17312 Fc-EMP fu
26	231	75.5	39	5	ABB73406	Abb73406 EPO mimet
27	203	66.3	145	7	ADJ73529	Adj73529 Erythropo
28	192	62.7	36	3	AAB17313	Aab17313 EMP-Fc fu
29	192	62.7	36	5	ABB73407	Abb73407 EPO mimet
30	168.5	55.1	253	3	AAB16964	Aab16964 Fc-EMP pr
31	168.5	55.1	253	5	ABB73415	Abb73415 Fc-EPO mi
32	161.5	52.8	253	3	AAB16965	Aab16965 EMP-Fc pr
33	161.5	52.8	253	5	ABB73416	Abb73416 EPO mimet
34	156	51.0	25	5	ABB73394	Abb73394 EPO-mimet
35	156	51.0	26	3	AAB17930	Aab17930 EPO-mimet
36	155.5	50.8	251	9	ADZ44485	Adz44485 Erythropo
37	152	49.7	51	8	ADJ52126	Adj52126 CH1 delet
38	152	49.7	51	8	ADJ52127	Adj52127 CH1 delet
39	152	49.7	269	8	ADJ52120	Adj52120 CH1 delet
40	151	49.3	266	8	ADJ52121	Adj52121 CH1 delet
41	150.5	49.2	37	8	ADJ52122	Adj52122 CH1 delet
42	148	48.4	131	7	ADJ73539	Adj73539 Erythropo
43	145.5	47.5	129	7	ADJ73537	Adj73537 Erythropo
44	144.5	47.2	249	9	ADZ44484	Adz44484 Erythropo
45	144.5	47.2	249	9	ADZ44490	Adz44490 Erythropo

ALIGNMENTS

RESULT 1
ABB73392
ID ABB73392 standard; peptide; 49 AA.
XX
AC ABB73392;
XX
DT 05-APR-2002 (first entry)
XX
DE EPO-mimetic peptide SEQ ID NO:339.
XX
KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200183525-A2.
XX
PD 08-NOV-2001.
XX
PF 02-MAY-2001; 2001WO-US014310.
XX
PR 03-MAY-2000; 2000US-00563286.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
XX
DR WPI; 2002-130313/17.
XX
PT Novel vehicle-peptide molecule or its multimers useful for treating
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT diabetic retinopathy, obesity, sleep disorders and infertility.
XX
PS Claim 16; Page 90; 176pp; English.
XX
CC The present invention describes a vehicle-peptide molecule (I) or its

CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive, cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological, antianaemic, anorectic, antiinfertility, haemostatic, dermatological and neuroprotective activities. (I) can be used as a therapeutic or prophylactic agent as well as for screening purposes. (I) is useful for diagnosing diseases characterised by dysfunction of their associated protein of interest, for identifying normal or abnormal proteins of interest, as a part of diagnostic kit to detect the presence of their proteins of interest in a biological sample. Additionally, (I) is useful for treating inflammatory and autoimmune diseases, tumour growth, cancer, rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders, infertility, and neurological degenerative diseases. (I), comprising EPO-mimetic compounds are useful for treating disorders characterised by low red blood cell levels such as anaemia. The TPO-mimetic comprising compounds are useful for treating conditions that involve an existing megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic tumour which result in thrombocytopaenia, systemic lupus erythematosus, and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777 represent amino acid and nucleic acid sequences used in the exemplification of the present invention

XX

SQ Sequence 49 AA;

Query Match 100.0%; Score 306; DB 5; Length 49;
Best Local Similarity 100.0%; Pred. No. 9.4e-25;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GGGGGGGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGG 49
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGGGGGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGG 49
```

RESULT 2
AAB17283
ID AAB17283 standard; peptide; 50 AA.
XX
AC AAB17283;
XX
DT 31-OCT-2000 (first entry)
XX
DE EPO-mimetic peptide sequence SEQ ID NO:339.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US025044.
XX
PR 23-OCT-1998; 98US-0105371P.
PR 22-OCT-1999; 99US-00428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and pharmacologically
PT active peptides, useful for treating cancer and autoimmune diseases.
XX
PS Claim 16; Page 314; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an Fc domain, pharmacologically active peptides, and linkers. Where (I) is: (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-(L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2, P3, and P4 = are each independently sequences of pharmacologically active peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b, c, d, e, and f = are each independently 0 or 1, provided that at least 1 of a and b is 1. The composition can have cytostatic, antiasthmatic, thrombolytic and immunosuppressive activities. DNAs, vectors and host cells from the present invention can be used for producing pharmaceutical compositions. The compositions are useful for treating cancer, asthma, thrombosis, or autoimmune diseases. The use of an Fc domain (rather than a Fab domain) can provide a longer half-life or incorporate functions such as Fc receptor binding, protein A binding, complement fixation, and possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to AAB18003 represent nucleotide and amino acid sequences used in the exemplification of the present invention

XX

SQ Sequence 50 AA;

Query Match 100.0%; Score 306; DB 3; Length 50;
Best Local Similarity 100.0%; Pred. No. 9.6e-25;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GGGGGGGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGG 49
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 GGGGGGGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGG 50
```

RESULT 3
AAB16967
ID AAB16967 standard; protein; 277 AA.
XX
AC AAB16967;
XX
DT 31-OCT-2000 (first entry)
XX
DE Fc-EMP-EMP protein sequence SEQ ID NO:22.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US025044.
XX
PR 23-OCT-1998; 98US-0105371P.
PR 22-OCT-1999; 99US-00428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
DR N-PSDB; AAA69451.
XX
PT Novel composition of matter comprising an Fc domain and pharmacologically
PT active peptides, useful for treating cancer and autoimmune diseases.
XX
PS Example 3; Page 201-202; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an

CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and
CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC AAB18003 represent nucleotide and amino acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 277 AA;

Query Match 100.0%; Score 306; DB 3; Length 277;
Best Local Similarity 100.0%; Pred. No. 5e-24;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGGGGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGG 49
Db 229 GGGGGGGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGG 277

RESULT 4
ABB73418
ID ABB73418 standard; protein; 277 AA.
XX
AC ABB73418;
XX
DT 05-APR-2002 (first entry)
XX
DE Fc-EMP-EMP nucleic acid SEQ ID NO:22.
XX
KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200183525-A2.
XX
PD 08-NOV-2001.
XX
PF 02-MAY-2001; 2001WO-US014310.
XX
PR 03-MAY-2000; 2000US-00563286.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
XX
DR WPI; 2002-130313/17.
DR N-PSDB; ABL35768.
XX
PT Novel vehicle-peptide molecule or its multimers useful for treating
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT diabetic retinopathy, obesity, sleep disorders and infertility.
XX
PS Claim 12; Fig 16; 176pp; English.
XX
CC The present invention describes a vehicle-peptide molecule (I) or its
CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
CC neuroprotective activities. (I) can be used as a therapeutic or
CC prophylactic agent as well as for screening purposes. (I) is useful for
CC diagnosing diseases characterised by dysfunction of their associated
CC protein of interest, for identifying normal or abnormal proteins of
CC interest, as a part of diagnostic kit to detect the presence of their
CC proteins of interest in a biological sample. Additionally, (I) is useful
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
CC infertility, and neurological degenerative diseases. (I), comprising EPO-
CC mimetic compounds are useful for treating disorders characterised by low
CC red blood cell levels such as anaemia. The TPO-mimetic comprising
CC compounds are useful for treating conditions that involve an existing
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC represent amino acid and nucleic acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 277 AA;

Query Match 100.0%; Score 306; DB 5; Length 277;
Best Local Similarity 100.0%; Pred. No. 5e-24;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGGGGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGG 49
Db 229 GGGGGGGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGG 277

RESULT 5
ABB73393
ID ABB73393 standard; peptide; 49 AA.
XX
AC ABB73393;
XX
DT 05-APR-2002 (first entry)
XX
DE EPO-mimetic peptide SEQ ID NO:340.
XX
KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200183525-A2.
XX
PD 08-NOV-2001.
XX
PF 02-MAY-2001; 2001WO-US014310.
XX
PR 03-MAY-2000; 2000US-00563286.
XX
PA (AMGE-) AMGEN INC.
XX

PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
XX
DR WPI; 2002-130313/17.
XX
PT Novel vehicle-peptide molecule or its multimers useful for treating
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT diabetic retinopathy, obesity, sleep disorders and infertility.
XX
PS Claim 16; Page 90; 176pp; English.
XX
CC The present invention describes a vehicle-peptide molecule (I) or its
CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
CC neuroprotective activities. (I) can be used as a therapeutic or
CC prophylactic agent as well as for screening purposes. (I) is useful for
CC diagnosing diseases characterised by dysfunction of their associated
CC protein of interest, for identifying normal or abnormal proteins of
CC interest, as a part of diagnostic kit to detect the presence of their
CC proteins of interest in a biological sample. Additionally, (I) is useful
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
CC infertility, and neurological degenerative diseases. (I), comprising EPO-
CC mimetic compounds are useful for treating disorders characterised by low
CC red blood cell levels such as anaemia. The TPO-mimetic comprising
CC compounds are useful for treating conditions that involve an existing
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC represent amino acid and nucleic acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 49 AA;

Query Match 90.2%; Score 276; DB 5; Length 49;
Best Local Similarity 100.0%; Pred. No. 1.3e-21;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 GGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGG 49
||||||||||||||||||||||||||||||||||||||||||||
Db 1 GGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGG 44

RESULT 6
AAB17284
ID AAB17284 standard; peptide; 50 AA.
XX
AC AAB17284;
XX
DT 31-OCT-2000 (first entry)
XX
DE EPO-mimetic peptide sequence SEQ ID NO:340.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US025044.
XX
PR 23-OCT-1998; 98US-0105371P.
PR 22-OCT-1999; 99US-00428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and pharmacologically
PT active peptides, useful for treating cancer and autoimmune diseases.
XX
PS Claim 16; Page 314; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and
CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC AAB18003 represent nucleotide and amino acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 50 AA;

Query Match 90.2%; Score 276; DB 3; Length 50;
Best Local Similarity 100.0%; Pred. No. 1.3e-21;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 GGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGG 49
||||||||||||||||||||||||||||||||||||||||||||
Db 1 GGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGG 44

RESULT 7
AAB17314
ID AAB17314 standard; peptide; 57 AA.
XX
AC AAB17314;
XX
DT 31-OCT-2000 (first entry)
XX
DE EMP-EMP-Fc fusion construction peptide SEQ ID NO:417.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US025044.
XX
PR 23-OCT-1998; 98US-0105371P.
PR 22-OCT-1999; 99US-00428082.
XX
PA (AMGE-) AMGEN INC.
XX

PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and pharmacologically
PT active peptides, useful for treating cancer and autoimmune diseases.
XX
PS Example 3; Page 342; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and
CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC AAB18003 represent nucleotide and amino acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 57 AA;

Query Match 90.2%; Score 276; DB 3; Length 57;
Best Local Similarity 100.0%; Pred. No. 1.5e-21;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 GGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGG 49
||||||||||||||||||||||||||||||||||||||||||||||
Db 2 GGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGG 45

RESULT 8
ABB73408
ID ABB73408 standard; peptide; 57 AA.
XX
AC ABB73408;
XX
DT 05-APR-2002 (first entry)
XX
DE EMP-EMP gene construction related peptide SEQ ID NO:417.
XX
KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200183525-A2.
XX
PD 08-NOV-2001.
XX
PF 02-MAY-2001; 2001WO-US014310.
XX
PR 03-MAY-2000; 2000US-00563286.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
XX
DR WPI; 2002-130313/17.
XX
PT Novel vehicle-peptide molecule or its multimers useful for treating
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT diabetic retinopathy, obesity, sleep disorders and infertility.
XX
PS Example 3; Page 116; 176pp; English.
XX
CC The present invention describes a vehicle-peptide molecule (I) or its
CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
CC neuroprotective activities. (I) can be used as a therapeutic or
CC prophylactic agent as well as for screening purposes. (I) is useful for
CC diagnosing diseases characterised by dysfunction of their associated
CC protein of interest, for identifying normal or abnormal proteins of
CC interest, as a part of diagnostic kit to detect the presence of their
CC proteins of interest in a biological sample. Additionally, (I) is useful
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
CC infertility, and neurological degenerative diseases. (I), comprising EPO-
CC mimetic compounds are useful for treating disorders characterised by low
CC red blood cell levels such as anaemia. The TPO-mimetic comprising
CC compounds are useful for treating conditions that involve an existing
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC represent amino acid and nucleic acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 57 AA;

Query Match 90.2%; Score 276; DB 5; Length 57;
Best Local Similarity 100.0%; Pred. No. 1.5e-21;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 GGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGG 49
||||||||||||||||||||||||||||||||||||||||||||||
Db 2 GGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGG 45

RESULT 9
AAB16966
ID AAB16966 standard; protein; 277 AA.
XX
AC AAB16966;
XX
DT 31-OCT-2000 (first entry)
XX
DE EMP-EMP-Fc protein sequence SEQ ID NO:20.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US025044.
XX

PR 23-OCT-1998; 98US-0105371P.
PR 22-OCT-1999; 99US-00428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
DR N-PSDB; AAA69450.
XX
PT Novel composition of matter comprising an Fc domain and pharmacologically
PT active peptides, useful for treating cancer and autoimmune diseases.
XX
PS Claim 17; Page 198-199; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and
CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC AAB18003 represent nucleotide and amino acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 277 AA;

Query Match 90.2%; Score 276; DB 3; Length 277;
Best Local Similarity 100.0%; Pred. No. 6.7e-21;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 GGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGG 49
||||||||||||||||||||||||||||||||||||||||||||
Db 2 GGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGG 45

RESULT 10
ABB73417
ID ABB73417 standard; protein; 278 AA.
XX
AC ABB73417;
XX
DT 05-APR-2002 (first entry)
XX
DE EMP-EMP-Fc amino acid SEQ ID NO:20.
XX
KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200183525-A2.
XX
PD 08-NOV-2001.
XX
PF 02-MAY-2001; 2001WO-US014310.
XX
PR 03-MAY-2000; 2000US-00563286.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
XX
DR WPI; 2002-130313/17.
DR N-PSDB; ABL35767.
XX
PT Novel vehicle-peptide molecule or its multimers useful for treating
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT diabetic retinopathy, obesity, sleep disorders and infertility.
XX
PS Claim 12; Fig 15; 176pp; English.
XX
CC The present invention describes a vehicle-peptide molecule (I) or its
CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
CC neuroprotective activities. (I) can be used as a therapeutic or
CC prophylactic agent as well as for screening purposes. (I) is useful for
CC diagnosing diseases characterised by dysfunction of their associated
CC protein of interest, for identifying normal or abnormal proteins of
CC interest, as a part of diagnostic kit to detect the presence of their
CC proteins of interest in a biological sample. Additionally, (I) is useful
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
CC infertility, and neurological degenerative diseases. (I), comprising EPO-
CC mimetic compounds are useful for treating disorders characterised by low
CC red blood cell levels such as anaemia. The TPO-mimetic comprising
CC compounds are useful for treating conditions that involve an existing
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC represent amino acid and nucleic acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 278 AA;

Query Match 90.2%; Score 276; DB 5; Length 278;
Best Local Similarity 100.0%; Pred. No. 6.7e-21;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 GGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGG 49
||||||||||||||||||||||||||||||||||||||||||||
Db 2 GGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGG 45

RESULT 11
ADJ72562
ID ADJ72562 standard; peptide; 70 AA.
XX
AC ADJ72562;
XX
DT 06-MAY-2004 (first entry)
XX
DE EPO mimetic peptide sequence SeqID 14.
XX
KW mimetic; CDR mimetibody; gene therapy; transgenic; immune;
KW cardiovascular; infectious; malignant; neurologic disease; anaemia;
KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;
KW erythropoietin; EPO.
XX
OS Synthetic.
XX
PN WO2003084477-A2.
XX
PD 16-OCT-2003.

XX
PF 24-MAR-2003; 2003WO-US009139.
XX
PR 29-MAR-2002; 2002US-0368791P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Scallon BJ, Ghrayeb J;
XX
DR WPI; 2003-804237/75.
XX
PT New CDR mimetibody comprising a portion of a heavy or light chain
PT variable region comprising human framework or ligand binding region,
PT useful for preparing a composition for treating e.g., immune,
PT cardiovascular or neurologic disease.
XX
PS Disclosure; SEQ ID NO 14; 97pp; English.
XX
CC This invention relates to novel mammalian CDR mimetibodies, specific
CC portions or variants thereof. Specifically, it refers to an antibody
CC fragment where a protein has been inserted into, or replaces a portion
CC of, one or more CDR regions, such that each CDR mimetibody comprises at
CC least one portion of a heavy chain or light chain variable region, which
CC itself comprises at least one human framework region and at least one
CC ligand binding region (LBR). The present invention describes human
CC mimetibodies, including modified immunoglobulins and cleavage products
CC that can be useful in gene therapy and the generation of transgenic
CC plants and animals. Furthermore, the CDR mimetibody is useful for
CC preparing compositions for modulating, treating or reducing the symptoms
CC of immune, cardiovascular, infectious, malignant and/ or neurologic
CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,
CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This
CC peptide sequence is a erythropoietin (EPO) mimetic peptide sequence used
CC to make a mimetibody of the invention.
XX
SQ Sequence 70 AA;

Query Match 84.6%; Score 259; DB 7; Length 70;
Best Local Similarity 85.7%; Pred. No. 1e-19;
Matches 42; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 GGGGGGGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGG 49
|| |||||||||||||||||||| ||||||||||||||||||||
Db 19 GGSSKGGTYSCHFGPLTWVCKPQGGSSKXGGTYSCHFGPLTWVCKPQGG 67

RESULT 12
AAB17040
ID AAB17040 standard; peptide; 47 AA.
XX
AC AAB17040;
XX
DT 31-OCT-2000 (first entry)
XX
DE EPO-mimetic peptide sequence SEQ ID NO:96.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US025044.
XX
PR 23-OCT-1998; 98US-0105371P.
XX
PR 22-OCT-1999; 99US-00428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and pharmacologically
PT active peptides, useful for treating cancer and autoimmune diseases.
XX
PS Claim 13; Page 228; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and
CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC AAB18003 represent nucleotide and amino acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 47 AA;

Query Match 81.4%; Score 249; DB 3; Length 47;
Best Local Similarity 90.9%; Pred. No. 7.8e-19;
Matches 40; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 6 GGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGG 49
|||||||||||||||||||| ||||||||||||||||||||
Db 1 GGTYSCHFGPLTWVCKPQGGSSKXGGTYSCHFGPLTWVCKPQGG 44

RESULT 13
ADJ52198
ID ADJ52198 standard; peptide; 47 AA.
XX
AC ADJ52198;
XX
DT 06-MAY-2004 (first entry)
XX
DE CH1 deleted mimetibody-related peptide SeqID14.
XX
KW CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;
KW fungicide; gene therapy; immune disorder; cardiovascular disease;
KW arrhythmia; hypertension; heart failure; neurodegenerative;
KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;
KW cancerous condition; infectious disease; bacterial infection;
KW viral infection; fungal infection.
XX
OS Unidentified.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 24
FT /label= OTHER
FT /note= "OTHER= linker"
XX
PN WO2004002417-A2.
XX
PD 08-JAN-2004.
XX

PF 27-JUN-2003; 2003WO-US020347.
XX
PR 28-JUN-2002; 2002US-0392431P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Ghrayeb J, Scallon BJ, Nesspor TC;
PI Kutoloski KA;
XX
DR WPI; 2004-082870/08.
XX
PT New CH1-deleted mimetibody polypeptides and nucleic acids, useful for
PT modulating, treating, alleviating, preventing an immune, cardiovascular,
PT or neurodegenerative disease or disorder, anemia, cancer, or infectious
PT diseases.
XX
PS Example 1; SEQ ID NO 14; 129pp; English.
XX
CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an immunosuppressive,
CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,
CC antibacterial, virucide or fungicide activity. In addition, the disclosed
CC sequences may prove useful for gene therapy. The CH1-deleted mimetibody
CC is useful for diagnosing or treating a disease condition in a cell,
CC tissue, organ or animal, specifically for modulating, treating,
CC alleviating, preventing the incidence or reducing the symptoms of an
CC immune, cardiovascular (for example arrhythmia, hypertension or heart
CC failure), or neurodegenerative (for example multiple sclerosis, dementia
CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous
CC conditions, or infectious diseases (for example bacterial, viral or
CC fungal infection). The present sequence is that of a peptide which may be
CC used during the creation of a mimetibody of the invention.
XX
SQ Sequence 47 AA;

Query Match 81.4%; Score 249; DB 8; Length 47;
Best Local Similarity 90.9%; Pred. No. 7.8e-19;
Matches 40; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 6 GGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGG 49
||||||||||||||||||||||||||||||||||||||||||||||
Db 1 GGTYSCHFGPLTWVCKPQGGSSKXGGTYSCHFGPLTWVCKPQGG 44

RESULT 14
ADJ51160
ID ADJ51160 standard; peptide; 47 AA.
XX
AC ADJ51160;
XX
DT 06-MAY-2004 (first entry)
XX
DE CH1 deleted mimetibody-related peptide SeqID14.
XX
KW CH1 deleted mimetibody; osteopathic; cardiovascular-Gen;
KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;
KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;
KW antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;
KW ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;
KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;
KW dental disorder; oral disorder; dermatological disorder; ear disorder;
KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;
KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;
KW obstetric disorder; haematologic disorder; immunological disorder;
KW allergic disorder; infectious disorder; musculoskeletal disorder;
KW oncological disorder; neurological disorder; nutritional disorder;
KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;
KW renal disorder; pulmonary disorder.
XX
OS Unidentified.
OS Synthetic.
XX

FH Key Location/Qualifiers
FT Misc-difference 24
FT /label= OTHER
FT /note= "OTHER= linker"
XX
PN WO2004002424-A2.
XX
PD 08-JAN-2004.
XX
PF 30-JUN-2003; 2003WO-US020495.
XX
PR 28-JUN-2002; 2002US-0392431P.
PR 19-SEP-2002; 2002US-0412144P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Ghrayeb J, Scallon BJ, Nesspor TC;
PI Kutoloski KA;
XX
DR WPI; 2004-082872/08.
XX
PT New CH1 deleted mimetibody polypeptide and nucleic acid, useful for
PT diagnosing, preventing or treating cardiovascular, dermatologic,
PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and
PT nutritional disorders.
XX
PS Claim 8; SEQ ID NO 14; 123pp; English.
XX
CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an osteopathic,
CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,
CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,
CC immunomodulator, antiallergic, muscular-Gen, cytostatic,
CC antiinflammatory, neuroleptic, ophthalmological, nephrotropic or
CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-
CC modulator or cytokine-agonist. The methods and compositions of the
CC present invention are useful for the diagnosis, prevention and/or
CC treatment of diseases or conditions associated with aberrant expression
CC or activity of the CH1 deleted mimetibody, such as a bone or joint,
CC cardiovascular, dental or oral, dermatological, ear, nose or throat,
CC endocrine, metabolic, gastrointestinal, gynaecological, hepatic,
CC obstetric, haematologic, immunological, allergic, infectious,
CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,
CC pediatric, psychiatric, renal or pulmonary disorders. The present
CC sequence is that of a peptide which may be used during the creation of a
CC mimetibody of the invention.
XX
SQ Sequence 47 AA;

Query Match 81.4%; Score 249; DB 8; Length 47;
Best Local Similarity 90.9%; Pred. No. 7.8e-19;
Matches 40; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 6 GGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGG 49
||||||||||||||||||||||||||||||||||||||||||||||
Db 1 GGTYSCHFGPLTWVCKPQGGSSKXGGTYSCHFGPLTWVCKPQGG 44

RESULT 15
AAB17036
ID AAB17036 standard; peptide; 40 AA.
XX
AC AAB17036;
XX
DT 31-OCT-2000 (first entry)
XX
DE EPO-mimetic peptide sequence SEQ ID NO:92.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;

KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US025044.
XX
PR 23-OCT-1998; 98US-0105371P.
PR 22-OCT-1999; 99US-00428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and pharmacologically
PT active peptides, useful for treating cancer and autoimmune diseases.
XX
PS Claim 13; Page 226; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and
CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC AAB18003 represent nucleotide and amino acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 40 AA;

Query Match 78.4%; Score 240; DB 3; Length 40;
Best Local Similarity 90.9%; Pred. No. 5.8e-18;
Matches 40; Conservative 0; Mismatches 0; Indels 4; Gaps 1;

```
Qy          6 GGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGG 49
              |||||||||||||||||    ||||||||||||||||||||||
Db          1 GGTYSCHFGPLTWVCKPQ----GGGGTYSCHFGPLTWVCKPQGG 40
```

Search completed: March 31, 2006, 16:22:27
Job time : 188.93 secs

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: March 31, 2006, 16:22:51 ; Search time 30.4726 Seconds
(without alignments)
154.717 Million cell updates/sec

Title: US-10-609-217-339
Perfect score: 306
Sequence: 1 GGGGGGGTYSCHFGPLTWVC.....GGTYSCHFGPLTWVCKPQGG 49

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	94.5	30.9	1585	2	T31611	hypothetical prote
2	90.5	29.6	280	2	G84839	late embryogenesis
3	89.5	29.2	214	1	KNNT2S	glycine-rich prote
4	88	28.8	209	2	JC4817	RNA-binding protei
5	88	28.8	2783	1	A41948	alpha-fetoprotein
6	86	28.1	299	2	T00837	glycine-rich prote
7	84.5	27.6	64	2	S53051	glycine rich prote
8	84.5	27.6	201	2	P84596	glycine-rich prote
9	84.5	27.6	307	2	T27609	hypothetical prote
10	84.5	27.6	388	2	T29173	hypothetical prote
11	84	27.5	299	2	T05494	glycine-rich prote
12	84	27.5	312	2	T25048	hypothetical prote
13	83.5	27.3	371	2	E88633	protein F56B3.1 [i
14	83	27.1	411	2	A49127	homeotic protein A
15	83	27.1	643	1	KRHU2	keratin 1, type II
16	83	27.1	910	2	A34721	androgen receptor
17	83	27.1	911	2	B34721	androgen receptor
18	83	27.1	919	2	A39248	androgen receptor
19	83	27.1	1381	2	E70806	hypothetical glyci
20	82.5	27.0	200	2	S10334	glycine-rich prote
21	82.5	27.0	290	2	T23416	hypothetical prote
22	82.5	27.0	316	2	T20497	hypothetical prote
23	82.5	27.0	405	2	T29167	hypothetical prote
24	82	26.8	207	2	B44994	eggshell protein 2
25	82	26.8	212	2	A44994	eggshell protein 1
26	81.5	26.6	255	2	B84777	hypothetical prote
27	81.5	26.6	482	2	T48337	hypothetical prote
28	80.5	26.3	203	1	JQ1061	glycine-rich prote
29	80	26.1	369	1	TVFVAF	transforming prote
30	80	26.1	481	2	A35628	loricrin - mouse
31	80	26.1	1084	2	T04103	sucrose-phosphate
32	79.5	26.0	325	2	G96718	unknown protein, 5
33	79	25.8	694	2	F70868	hypothetical glyci
34	78.5	25.7	188	2	S49192	GCR 1 protein - fr
35	78.5	25.7	291	1	S31415	glycine-rich prote
36	78.5	25.7	404	2	S54729	RNA-binding protei
37	78	25.5	207	2	T07381	glycine-rich prote
38	78	25.5	239	2	S49193	GCR 101 protein -
39	78	25.5	343	2	T29547	hypothetical prote
40	77.5	25.3	165	1	KNRZG1	glycine-rich cell
41	77.5	25.3	183	2	PN0109	keratin-like prote
42	77.5	25.3	256	2	T03371	glycine-rich prote
43	77.5	25.3	2174	2	E95965	hypothetical glyci
44	77	25.2	196	2	S49194	GCR 17 protein - f
45	77	25.2	263	2	A34466	calpain (EC 3.4.22

ALIGNMENTS

RESULT 1
T31611
hypothetical protein Y50E8A.g - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C;Accession: T31611
R;Steward, C.
submitted to the EMBL Data Library, September 1999
A;Reference number: Z21047
A;Accession: T31611
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1585 <WIL>
A;Cross-references: UNIPARC:UPI000017BC9F; EMBL:AL117200; NID:e1549770; PIDN:CAB55050.1;
A;Experimental source: clone Y50E8A
C;Genetics:
A;Gene: CESP:Y50E8A.g
A;Introns: 25/3; 60/1; 133/2; 217/3; 270/3; 337/2; 400/1; 746/2

Query Match 30.9%; Score 94.5; DB 2; Length 1585;
Best Local Similarity 42.9%; Pred. No. 0.07;
Matches 21; Conservative 1; Mismatches 6; Indels 21; Gaps 2;

```
Qy          1 GGGGGGGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGG 49
              |||| || |:          || |||||||||            || ||
Db        463 GGGGAGGGYA----------KPSGGGGGGGG----------YAKPSGG 490
```

RESULT 2
G84839
late embryogenesis abundant M17 protein [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: G84839
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: G84839
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-280 <STO>
A;Cross-references: UNIPROT:Q9S7S3; UNIPARC:UPI000009DE6F; GB:AE002093; NID:g3894196; PI
C;Genetics:
A;Gene: At2g41260
A;Map position: 2

Query Match 29.6%; Score 90.5; DB 2; Length 280;
Best Local Similarity 30.4%; Pred. No. 0.037;
Matches 21; Conservative 4; Mismatches 13; Indels 31; Gaps 2;

```
Qy          1 GGGGGGGTYSCHF--------GPLTWVCKPQ----------------------GGGGGG 29
              |||||||   | :        |   : |: |                        ||||||
Db        123 GGGGGGGRGGCRWGCCGGWWRGRCRYCCRSQAEASEVVETVEPNDVEPQQGGRGGGGGGG 182

Qy         30 GGTYSCHFG 38
              ||   | :|
Db        183 GGRGGCRWG 191
```

RESULT 3
KNNT2S
glycine-rich protein 2 - wood tobacco
C;Species: Nicotiana sylvestris (wood tobacco)
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
C;Accession: S17731
R;Obokata, J.; Ohme, M.; Hayashida, N.
Plant Mol. Biol. 17, 953-955, 1991
A;Title: Nucleotide sequence of a cDNA clone encoding a putative glycine-rich protein of
A;Reference number: S17731; MUID:92003709; PMID:1912512
A;Accession: S17731
A;Molecule type: mRNA
A;Residues: 1-214 <OBO>
A;Cross-references: UNIPROT:P27484; UNIPARC:UPI000012BB03; EMBL:X60007; NID:g19742; PIDN
C;Superfamily: Arabidopsis glycine-rich protein 2; cold shock domain homology
C;Keywords: zinc finger
F;11-71/Domain: cold shock domain homology <CSD>
F;82-158/Region: glycine-rich
F;159-172/Region: zinc finger CCHC motif
F;176-195/Region: glycine-rich
F;196-209/Region: zinc finger CCHC motif

  Query Match   29.2%; Score 89.5; DB 1; Length 214;
  Best Local Similarity   51.2%; Pred. No. 0.037;
  Matches   21; Conservative   0; Mismatches   5; Indels   15; Gaps   3;

```
Qy          1 GGGGGGGTYSCHFGPLTWVCKPQGGGGGGGGGTYSC----HF 37
              |||||||      ||          |||||||| | |     ||
Db        176 GGGGGGG----RFG-------GGGGGGGGGGCYKCGEDGHF 205
```

RESULT 4
JC4817
RNA-binding protein RZ-1 - wood tobacco
C;Species: Nicotiana sylvestris (wood tobacco)
C;Date: 15-Aug-1996 #sequence_revision 15-Oct-1996 #text_change 09-Jul-2004
C;Accession: JC4817; PC4175
R;Hanano, S.; Sugita, M.; Sugiura, M.
DNA Res. 3, 65-71, 1996
A;Title: Structure and expression of the tobacco nuclear gene encoding RNA-binding prote
A;Reference number: JC4817; MUID:96397973; PMID:8804857
A;Accession: JC4817
A;Molecule type: mRNA
A;Residues: 1-209 <HAN1>
A;Cross-references: UNIPROT:Q42412; UNIPARC:UPI00000ACA12; DDBJ:D83696; NID:g1395192; PI
A;Accession: PC4175
A;Molecule type: protein
A;Residues: 1-209 <HAN2>
A;Cross-references: UNIPARC:UPI00000ACA12
A;Experimental source: leaf
C;Comment: This protein, localizes in the nucleus, contains a zinc finger motif and a co
C;Genetics:
A;Gene: rz-1
A;Introns: 36/3
C;Superfamily: glycine-rich RNA-binding protein; ribonucleoprotein repeat homology
F;7-74/Domain: ribonucleoprotein repeat homology <RRM1>

  Query Match   28.8%; Score 88; DB 2; Length 209;
  Best Local Similarity   41.7%; Pred. No. 0.051;
  Matches   20; Conservative   5; Mismatches   11; Indels   12; Gaps   3;

```
Qy          1 GGGGGGGTYSC----HFGPLTWVCKPQGG-----GGGGGGTYSCHFGP 39
              | ||||  ::|    ||      |  :||     ||||||: |  :||
Db        122 GSGGGGDCFNCGKPGHFAR---ECPSEGGRGGRYGGGGGGSRSSGYGP 166
```

RESULT 5
A41948
alpha-fetoprotein enhancer-binding protein - human
N;Alternate names: ATBF1 protein
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 15-Oct-1999
C;Accession: A41948
R;Morinaga, T.; Yasuda, H.; Hashimoto, T.; Higashio, K.; Tamaoki, T.
Mol. Cell. Biol. 11, 6041-6049, 1991
A;Title: A human alpha-fetoprotein enhancer-binding protein, ATBF1, contains four homeodc
A;Reference number: A41948; MUID:92049333; PMID:1719379
A;Accession: A41948
A;Molecule type: mRNA
A;Residues: 1-2783 <MOR>
A;Cross-references: UNIPARC:UPI0000156381; GB:D10250; GB:D90395; NID:g219429; PIDN:BAA01(
A;Note: sequence extracted from NCBI backbone (NCBIN:66271, NCBIP:66276)
C;Genetics:
A;Gene: GDB:ATBF1
A;Cross-references: GDB:392090; OMIM:104155
A;Map position: 16q22.3-16q23.1
C;Superfamily: alpha-fetoprotein enhancer-binding protein; homeobox homology
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation; zinc finger
F;72-94/Region: zinc finger CCHH motif
F;128-150/Region: zinc finger CCHH motif
F;176-198/Region: zinc finger CCHH motif
F;311-332/Region: zinc finger CCHH motif
F;340-361/Region: zinc finger CCHH motif
F;448-471/Region: zinc finger CCHH motif
F;489-509/Region: zinc finger CCHH motif
F;517-538/Region: zinc finger CCHH motif
F;633-655/Region: zinc finger CCHH motif
F;684-706/Region: zinc finger CCHH motif
F;719-773/Region: serine/threonine-rich
F;809-958/Region: glutamine-rich
F;1071-1092/Region: zinc finger CCHH motif
F;1117-1211/Region: proline-rich
F;1232-1288/Domain: homeobox homology <HOX1>
F;1329-1385/Domain: homeobox homology <HOX2>
F;1416-1437/Region: zinc finger CCHC motif
F;1618-1638/Region: zinc finger CCHH motif
F;1728-1784/Domain: homeobox homology <HOX3>
F;1799-1820/Region: zinc finger CCHH motif
F;2033-2089/Domain: homeobox homology <HOX4>
F;2112-2134/Region: zinc finger CCHH motif
F;2545-2566/Region: zinc finger CCHC motif
F;2585-2607/Region: glycine-rich
F;2611-2633/Region: zinc finger CCHH motif
F;2650-2737/Region: serine/threonine-rich

  Query Match   28.8%; Score 88; DB 1; Length 2783;
  Best Local Similarity   48.6%; Pred. No. 0.5;
  Matches   17; Conservative   1; Mismatches   1; Indels   16; Gaps   1;

```
Qy          1 GGGGGGGTYSCHFGPLTWVCKPQGGGGGGGGGTYSC 35
              |||||||                |||||||:| |
Db       2593 GGGGGGG---------------GGGGGGGGSYHC 2611
```

RESULT 6
T00837
glycine-rich protein T13L16.11 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 31-Dec-2004
C;Accession: T00837; D84557
R;de la Bastide, M.; Hameed, A.; Gnoj, L.; Jensen, K.; Shohdy, N.; Gottesman, T.; Haberma
 McCombie, W.R.
submitted to the EMBL Data Library, January 1999
A;Description: A. thaliana BAC T13L16 from chromosome IV, top arm.
A;Reference number: Z14205
A;Accession: T00837

```
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-299 <DEL>
A;Cross-references: UNIPROT:Q94C69; UNIPARC:UPI0000177E58; EMBL:AC003952; NID:g2708736;
A;Experimental source: cultivar Columbia
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: D84557
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-299 <STO>
A;Cross-references: UNIPARC:UPI0000177E58; GB:AE002093; NID:g2708747; PIDN:AAD03571.1; G
C;Genetics:
A;Gene: At2g17870; T13L16.11
A;Map position: 2
C;Superfamily: cold shock domain homology
F;11-71/Domain: cold shock domain homology <CSD>

  Query Match             28.1%;  Score 86;  DB 2;  Length 299;
  Best Local Similarity   40.9%;  Pred. No. 0.11;
  Matches   27;  Conservative    2;  Mismatches   17;  Indels   20;  Gaps    6;

Qy          1 GGGGGGG--TYSC-HFGPLTWVCK------PQGGGG----GGGGTYSC----HFGPLTWV 43
              |||||||   |||   | |   |:        ||||    || | | |    ||
Db        152 GGGGGGGRPCYSCGEVGHLAKDCRGGSGGNRYGGGGGRGSGGDGCYMCGGVGHFAR---D 208

Qy         44 CKPQGG 49
              |:  ||
Db        209 CRQNGG 214


RESULT 7
S53051
glycine rich protein - barley (fragment)
C;Species: Hordeum vulgare (barley)
C;Date: 08-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 09-Jul-2004
C;Accession: S53051
R;Molina, A.; Mena, M.; Garcia-Olmedo, F.; Carbonero, P.
submitted to the EMBL Data Library, March 1995
A;Description: Developmental and pathogen-induced expression of two barley genes encodin
A;Reference number: S53050
A;Accession: S53051
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-64 <MOL>
A;Cross-references: UNIPROT:Q40051; UNIPARC:UPI00000AA403; EMBL:Z48625; NID:g728595; PID
C;Superfamily: Arabidopsis glycine-rich protein 3

  Query Match             27.6%;  Score 84.5;  DB 2;  Length 64;
  Best Local Similarity   46.3%;  Pred. No. 0.039;
  Matches   19;  Conservative    2;  Mismatches    7;  Indels   13;  Gaps    2;

Qy          2 GGGGGGTYSCHFGPLTWVCKPQGGGGG----GGGTYSCHFG 38
              || ||| |  | |          |||||    |||:  ||:|
Db          4 GGKGGGGYPGHGG---------GGGGGYPGHGGGSSGCHWG 35


RESULT 8
F84596
glycine-rich protein (AtGRP2) [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: F84596
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: F84596
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-201 <STO>
A;Cross-references: UNIPROT:Q38896; UNIPARC:UPI0000000E38; GB:AE002093; NID:g4803937; PID
C;Genetics:
A;Gene: At2g21060
A;Map position: 2
C;Superfamily: Arabidopsis glycine-rich protein 2; cold shock domain homology

  Query Match             27.6%;  Score 84.5;  DB 2;  Length 201;
  Best Local Similarity   39.6%;  Pred. No. 0.11;
  Matches   21;  Conservative    4;  Mismatches   19;  Indels    9;  Gaps    2;

Qy          1 GGGGGGGTYSCHFGPLTWVCKPQGGGGGGGGTYSC----HFGPLTWVCKPQGG 49
              ||| |||:|   :|      :  || |||||  ||      | :   |   ||
Db        109 GGGRGGGSYGGGYGG-----RGSGGRGGGGGDNSCFKCGEPGHMARECSQGGG 156


RESULT 9
T27609
hypothetical protein ZC477.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C;Accession: T27609
R;Du, Z.
submitted to the EMBL Data Library, November 1995
A;Description: The sequence of C. elegans cosmid ZC477.
A;Reference number: Z20392
A;Accession: T27609
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-307 <DUZ>
A;Cross-references: UNIPARC:UPI0000177E53; EMBL:U40802; PIDN:AAA81510.1; CESP:ZC477.1
C;Genetics:
A;Gene: CESP:ZC477.1
A;Introns: 32/1; 275/1
C;Superfamily: Phaseolus glycine-rich cell wall protein 1.8

  Query Match             27.6%;  Score 84.5;  DB 2;  Length 307;
  Best Local Similarity   47.4%;  Pred. No. 0.16;
  Matches   18;  Conservative    3;  Mismatches   12;  Indels    5;  Gaps    1;

Qy          1 GGGGGGGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFG 38
              |||||||  | :||  :        |||| |  | :||
Db        226 GGGGGGGATSAYFGVGSGAM-----GGGGAGAQSAYFG 258


RESULT 10
T29173
hypothetical protein T28H11.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T29173
R;Nelson, J.; Wohldmann, P.
submitted to the EMBL Data Library, July 1996
A;Description: The sequence of C. elegans cosmid T28H11.
A;Reference number: Z20582
A;Accession: T29173
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-388 <NEL>
A;Cross-references: UNIPROT:Q23062; UNIPARC:UPI0000177E52; EMBL:U64609; PIDN:AAB04604.1;
A;Experimental source: strain Bristol N2; clone T28H11
C;Genetics:
A;Gene: CESP:T28H11.1
A;Map position: 4
A;Introns: 354/1
C;Superfamily: Phaseolus glycine-rich cell wall protein 1.8

  Query Match             27.6%;  Score 84.5;  DB 2;  Length 388;
  Best Local Similarity   47.4%;  Pred. No. 0.19;
  Matches   18;  Conservative    3;  Mismatches   12;  Indels    5;  Gaps    1;
```

```
Qy          1 GGGGGGGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFG 38
              |||||||  | :||  :        ||||  |  | :||
Db        305 GGGGGGGATSAYFGVGSGAM-----GGGGAGAQSAYFG 337
```

RESULT 11
T05494
glycine-rich protein T19K4.150 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C;Accession: T05494
R;Bevan, M.; Wedler, H.; Wambutt, R.; Hoheisel, J.; Mewes, H.W.; Mayer, K.F.X.; Schuelle
submitted to the Protein Sequence Database, April 1998
A;Reference number: Z15418
A;Accession: T05494
A;Molecule type: DNA
A;Residues: 1-299 <BEV>
A;Cross-references: UNIPROT:O65639; UNIPARC:UPI00000A694F; EMBL:AL022373
A;Experimental source: cultivar Columbia; BAC clone T19K4
C;Genetics:
A;Map position: 4
A;Note: T19K4.150
C;Superfamily: Arabidopsis glycine-rich protein 2; cold shock domain homology
F;13-73/Domain: cold shock domain homology <CSD>

  Query Match 27.5%; Score 84; DB 2; Length 299;
  Best Local Similarity 44.0%; Pred. No. 0.17;
  Matches 22; Conservative 3; Mismatches 9; Indels 16; Gaps 4;

```
Qy          2 GGGGGGTYSCHFGPLTWVCKP--QGGGGGGG--------GTYSC----HF 37
              ||||  | |:|  | |  : |    |||||||        | |:|    ||
Db         96 GGGGSGCYNC--GELGHISKDCGIGGGGGGGERRSRGGEGCYNCGDTGHF 143
```

RESULT 12
T25048
hypothetical protein T21B4.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T25048
R;Smye, R.
submitted to the EMBL Data Library, October 1996
A;Reference number: Z19974
A;Accession: T25048
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-312 <WIL>
A;Cross-references: UNIPROT:O18097; UNIPARC:UPI0000081E7D; EMBL:Z81124; PIDN:CAB03369.1;
A;Experimental source: clone T21B4
C;Genetics:
A;Gene: CESP:T21B4.2
A;Map position: 2
A;Introns: 21/1; 53/3

  Query Match 27.5%; Score 84; DB 2; Length 312;
  Best Local Similarity 48.6%; Pred. No. 0.18;
  Matches 18; Conservative 1; Mismatches 8; Indels 10; Gaps 1;

```
Qy          2 GGGGGGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFG 38
              ||||||  |:          ||||||||  ||  |
Db         99 GGGGGGGYA----------TGGGGGGGGGGCCSCGIG 125
```

RESULT 13
E88633
protein F56B3.1 [imported] - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C;Accession: E88633
R;anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
A;Reference number: A75000; MUID:99069613; PMID:9851916
A;Note: see websites genome.wustl.edu/gsc/C_elegans/ and www_sanger.ac.uk/Projects/C_eleg
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A;Accession: E88633
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-371 <STO>
A;Cross-references: UNIPROT:O45114; UNIPARC:UPI00000772FC; GB:chr_IV; PIDN:AAC02612.1; PI
A;Note: contains similarity to collagens
C;Genetics:
A;Gene: F56B3.1
A;Map position: 4

  Query Match 27.3%; Score 83.5; DB 2; Length 371;
  Best Local Similarity 40.4%; Pred. No. 0.23;
  Matches 21; Conservative 1; Mismatches 17; Indels 13; Gaps 2;

```
Qy          1 GGGGGGGTYSCHFGPLTWVCKPQGG------GGGGGGTYSCHFGPLTWVCKP 46
              |||||||    | |          ||      ||||||    |    | :  | |
Db         98 GGGGGGGYGGGHGG-------GHGGAVGGGYGGGGGGGGGCQCSPSSNTCPP 142
```

RESULT 14
A49127
homeotic protein AmphiHox3 - Florida lancelet
C;Species: Branchiostoma floridae (Florida lancelet)
C;Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 31-Dec-2004
C;Accession: A49127; S24762; S33565
R;Holland, P.W.H.; Holland, L.Z.; Williams, N.A.; Holland, N.D.
Development 116, 653-661, 1992
A;Title: An amphioxus homeobox gene: sequence conservation, spatial expression during dev
A;Reference number: A49127; MUID:93170170; PMID:1363226
A;Accession: A49127
A;Molecule type: DNA
A;Residues: 1-411 <HOL>
A;Cross-references: UNIPROT:P50901; UNIPARC:UPI000012CAE8; EMBL:X68045; NID:g5719; PIDN:C
A;Note: sequence extracted from NCBI backbone (NCBIP:125689)
C;Genetics:
A;Gene: AmphiHox3
A;Introns: 119/1
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
F;136-192/Domain: homeobox homology <HOX>

  Query Match 27.1%; Score 83; DB 2; Length 411;
  Best Local Similarity 50.0%; Pred. No. 0.29;
  Matches 19; Conservative 1; Mismatches 14; Indels 4; Gaps 1;

```
Qy          1 GGGGGGGTYSCH----FGPLTWVCKPQGGGGGGGGTYS 34
              |||||||:|       |||   | ||    ||| |  |
Db         32 GGGGGGGSYGAEVTQDFGPSCSVRKPGSTTGGGTGNVS 69
```

RESULT 15
KRHU2
keratin 1, type II, cytoskeletal - human
N;Alternate names: 67K type II epidermal keratin; cytokeratin 1
C;Species: Homo sapiens (man)
C;Date: 04-Dec-1986 #sequence_revision 22-Oct-1999 #text_change 10-Dec-1999
C;Accession: A22940; A02950; A43342
R;Johnson, L.D.; Idler, W.W.; Zhou, X.M.; Roop, D.R.; Steinert, P.M.
Proc. Natl. Acad. Sci. U.S.A. 82, 1896-1900, 1985
A;Reference number: A22940; MUID:85166239; PMID:2580302
A;Accession: A22940
A;Molecule type: DNA
A;Residues: 1-643 <JOH>
A;Cross-references: UNIPARC:UPI0000173D59; GB:M98776; GB:M11215; GB:M11845; GB:M11846; NI
A;Note: translation of initiator Met is not shown
R;Steinert, P.M.; Parry, D.A.D.; Idler, W.W.; Johnson, L.D.; Steven, A.C.; Roop, D.R.
J. Biol. Chem. 260, 7142-7149, 1985
A;Title: Amino acid sequences of mouse and human epidermal type II keratins of M-r 67,000
iate filament subunits.
A;Reference number: A92535; MUID:85207740; PMID:2581964
A;Accession: A02950

```
A;Molecule type: mRNA
A;Residues: 151-183,'K',185-199,'M',201-204,'K',206-236,'S',238-239,'R',241-356,'Y',358-
,'S',638-643 <STE>
A;Cross-references: UNIPARC:UPI000016ABD0; GB:M10938; NID:g186787; PIDN:AAA36153.1; PID:
A;Experimental source: tissue neonatal foreskin
A;Note: the authors translated the codon CUG for residue 476 as Met
R;Chipev, C.C.; Korge, B.P.; Markova, N.; Bale, S.J.; DiGiovanna, J.J.; Compton, J.G.; S
Cell 70, 821-828, 1992
A;Title: A leucine----proline mutation in the H1 subdomain of keratin 1 causes epidermol
A;Reference number: A43342; MUID:92386601; PMID:1381288
A;Accession: A43342
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 144-146,'P',148-159,'P',161-183,'K',185-186 <CHI>
A;Cross-references: UNIPARC:UPI0000173D5A; GB:M98776; GB:M11215; GB:M11845; GB:M11846; N
A;Note: sequence extracted from NCBI backbone (NCBIP:112784)
C;Comment: The cytoskeletal and microfibrillar keratins are classified into two types, t
atin IF protein subunit appears to be a heterotetramer of two type I and two type II pro
C;Comment: Keratin 1 is expressed in terminally differentiating epidermis.
C;Genetics:
A;Gene: GDB:KRT1
A;Cross-references: GDB:128198; OMIM:139350
A;Map position: 12q11-12q13
A;Note: defects in this gene may result in epidermolytic hyperkeratosis
C;Complex: heterotetramer of two type I, usually keratin 10 (see PIR:KRHU0), and two typ
C;Superfamily: cytoskeletal keratin
C;Keywords: coiled coil; heterotetramer; intermediate filament
F;4-179/Domain: head <HED>
F;4-143/Region: E1 and V1 subdomains
F;14-179/Region: H1 subdomain
F;180-492/Domain: rod <ROD>
F;180-214/Region: coil 1A
F;215-226/Region: linker 1
F;227-327/Region: coil 1B
F;328-344/Region: linker 12
F;345-363/Region: coil 2A
F;364-371/Region: linker 2
F;372-492/Region: coil 2B
F;430/Region: stutter
F;493-643/Domain: tail <END>
F;493-512/Region: H2 subdomain
F;513-643/Region: V2 and E2 subdomains

  Query Match             27.1%;  Score 83;  DB 1;  Length 643;
  Best Local Similarity   42.6%;  Pred. No. 0.43;
  Matches   23;  Conservative    1;  Mismatches   16;  Indels   14;  Gaps    3;

Qy          1 GGGGGGGTYSCHFGPLTWVCKPQGGGG------GGGGTYSCHFGPLTWVCKPQG 48
              || ||||     ||         ||||      |||| |   :||   || | |
Db        105 GGFGGGGIGGGGFGGF-----GSGGGGFGGGGFGGGGGYGGGYGP---VCSPGG 150


Search completed: March 31, 2006, 16:37:21
Job time : 31.4726 secs
```

This Page Blank (uspto)

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: March 31, 2006, 16:09:36 ; Search time 183.567 Seconds
(without alignments)
188.328 Million cell updates/sec

Title: US-10-609-217-339
Perfect score: 306
Sequence: 1 GGGGGGGTYSCHFGPLTWVC.....GGTYSCHFGPLTWVCKPQGG 49

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	120	39.2	779	2	Q4R1V0_9CRUS	Q4r1v0 daphnia mag
2	119	38.9	775	2	Q4W7T9_9CRUS	Q4w7t9 daphnia mag
3	97	31.7	205	2	Q75QN9_WHEAT	Q75qn9 triticum ae
4	97	31.7	229	2	Q8LPA7_WHEAT	Q8lpa7 triticum ae
5	95.5	31.2	241	2	Q6YUR8_ORYSA	Q6yur8 oryza sativ
6	95	31.0	231	2	Q75QN8_WHEAT	Q75qn8 triticum ae
7	94	30.7	473	2	Q84WQ1_ARATH	Q84wq1 arabidopsis
8	93	30.4	249	2	Q7XGH3_ORYSA	Q7xgh3 oryza sativ
9	93	30.4	249	2	Q8S748_ORYSA	Q8s748 oryza sativ
10	92.5	30.2	318	2	Q61PL7_CAEBR	Q61pl7 caenorhabdi
11	90.5	29.6	197	2	Q84UR8_ORYSA	Q84ur8 oryza sativ
12	90.5	29.6	225	2	Q8L7T1_ARATH	Q8l7t1 arabidopsis
13	90.5	29.6	280	2	Q9S7S3_ARATH	Q9s7s3 arabidopsis
14	90	29.4	117	2	Q9VD49_DROME	Q9vd49 drosophila
15	89.5	29.2	214	1	GRP2_NICSY	P27484 nicotiana s
16	88.5	28.9	650	2	O97344_9BILA	O97344 sagitta ele
17	88	28.8	209	2	Q42412_NICSY	Q42412 nicotiana s
18	88	28.8	753	2	Q61BJ9_CAEBR	Q61bj9 caenorhabdi
19	87.5	28.6	770	2	Q9GNP1_CIOSA	Q9gnp1 ciona savig
20	87	28.4	305	2	Q9CX86_MOUSE	Q9cx86 mus musculu
21	87	28.4	324	2	O35295_MOUSE	O35295 mus musculu
22	87	28.4	494	2	Q9N3E0_CAEEL	Q9n3e0 caenorhabdi
23	86	28.1	299	2	Q7XJR8_ARATH	Q7xjr8 arabidopsis
24	86	28.1	301	2	Q94C69_ARATH	Q94c69 arabidopsis
25	86	28.1	342	2	Q9VKR8_DROME	Q9vkr8 drosophila
26	86	28.1	356	2	Q6NNY8_DROME	Q6nny8 drosophila
27	86	28.1	472	2	Q7YXK1_ASCSU	Q7yxk1 ascaris suu
28	86	28.1	712	2	Q4JF01_PLADU	Q4jf01 platynereis
29	85.5	27.9	410	2	Q6YUX6_ORYSA	Q6yux6 oryza sativ
30	85.5	27.9	414	2	Q5MNI9_DROSE	Q5mni9 drosophila
31	85.5	27.9	414	2	Q5MNJ0_DROMA	Q5mnj0 drosophila
32	85.5	27.9	414	2	Q5MNJ1_DROSI	Q5mnj1 drosophila
33	85.5	27.9	428	1	UNC4_DROME	O77215 drosophila
34	85	27.8	332	2	Q640X7_XENLA	Q640x7 xenopus lae
35	85	27.8	369	2	Q61HI5_CAEBR	Q61hi5 caenorhabdi
36	85	27.8	496	2	Q8JFQ5_ONCMY	Q8jfq5 oncorhynchu
37	84.5	27.6	64	2	Q40051_HORVU	Q40051 hordeum vul
38	84.5	27.6	93	2	Q54B94_DICDI	Q54b94 dictyosteli
39	84.5	27.6	167	2	Q9LQZ8_ARATH	Q9lqz8 arabidopsis
40	84.5	27.6	201	1	GRP2B_ARATH	Q38896 arabidopsis
41	84.5	27.6	201	2	Q5BIT2_ARATH	Q5bit2 arabidopsis
42	84.5	27.6	265	2	Q23347_CAEEL	Q23347 caenorhabdi
43	84.5	27.6	373	2	Q60TM5_CAEBR	Q60tm5 caenorhabdi
44	84.5	27.6	373	2	Q23062_CAEEL	Q23062 caenorhabdi
45	84.5	27.6	399	1	CAZ_DROME	Q27294 drosophila

ALIGNMENTS

RESULT 1
Q4R1V0_9CRUS
ID Q4R1V0_9CRUS PRELIMINARY; PRT; 779 AA.
AC Q4R1V0;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE VASA RNA helicase.
GN Name=vasa;
OS Daphnia magna.
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Branchiopoda; Diplostraca;
OC Cladocera; Anomopoda; Daphniidae; Daphnia.
OX NCBI_TaxID=35525;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sagawa K., Yamagata H., Shiga Y.;
RT "Exploring embryonic germ line development in the water flea, Daphnia
RT magna, by zinc-finger-containing VASA as a marker.";
RL Gene Expr. Patterns 5:669-678(2005).
DR EMBL; AB193327; BAE00180.1; -; Genomic_DNA.
KW Helicase.
SQ SEQUENCE 779 AA; 82342 MW; B6C30D45AA8B352F CRC64;

Query Match 39.2%; Score 120; DB 2; Length 779;
Best Local Similarity 49.1%; Pred. No. 0.00065;
Matches 28; Conservative 6; Mismatches 13; Indels 10; Gaps 4;

```
Qy          1 GGGGGGGTYSCH----FGPLTWVCKPQGGGGGGGGTYSCH----FGPLTWVCKPQGG 49
              |||||||: :||     |  :  | ||||||| |  :||     | ::  | ||||
Db        214 GGGGGGGSRACHKCGEEGHFSREC-PQGGGGGGSGPRTCHKCGEEGHMSRDC-PQGG 268
```

RESULT 2
Q4W7T9_9CRUS
ID Q4W7T9_9CRUS PRELIMINARY; PRT; 775 AA.
AC Q4W7T9;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE VASA RNA helicase.
GN Name=vasa;
OS Daphnia magna.
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Branchiopoda; Diplostraca;
OC Cladocera; Anomopoda; Daphniidae; Daphnia.
OX NCBI_TaxID=35525;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sagawa K., Yamagata H., Shiga Y.;
RT "Exploring embryonic germ line development in the water flea, Daphnia
RT magna, by zinc-finger-containing VASA as a marker.";
RL Gene Expr. Patterns 5:669-678(2005).
DR EMBL; AB193324; BAD99522.1; -; mRNA.
KW Helicase.
SQ SEQUENCE 775 AA; 82164 MW; E388E608BA098125 CRC64;

```
  Query Match             38.9%;  Score 119;  DB 2;  Length 775;
  Best Local Similarity   49.1%;  Pred. No. 0.00082;
  Matches   28;  Conservative    6;  Mismatches   13;  Indels   10;  Gaps    4;

Qy          1 GGGGGGGTYSCH----FGPLTWVCKPQGGGGGGGGTYSCH----FGPLTWVCKPQGG 49
              |||||||: :||      |  :  | |||||||| |  :||     | ::  | ||||
Db        210 GGGGGGGSRACHKCGEEGHFSREC-PQGGGGGGSGPRTCHKCGEEGHVSRDC-PQGG 264


RESULT 3
Q75QN9_WHEAT
ID   Q75QN9_WHEAT PRELIMINARY;      PRT;   205 AA.
AC   Q75QN9;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE   Cold shock domain protein 2 (Putative glycine-rich protein).
GN   Name=WCSP2; Synonyms=GRP;
OS   Triticum aestivum (Wheat).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC   Triticeae; Triticum.
OX   NCBI_TaxID=4565;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Nakaminami K., Imai R.;
RL   Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Seed;
RA   Nakamura T., Saito M., Vrinten P., Shimbata T.;
RL   Submitted (DEC-2003) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Contains 1 CSD (cold-shock) domain.
DR   EMBL; AB161682; BAD08700.1; -; mRNA.
DR   EMBL; AB158409; BAD06324.1; -; mRNA.
DR   Gramene; Q75QN9; -.
DR   GO; GO:0003677; F:DNA binding; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR011129; CSP.
DR   InterPro; IPR002059; CSP_DNA_bd.
DR   InterPro; IPR012340; OB_NA_bd_sub.
DR   InterPro; IPR001878; Znf_CCHC.
DR   Pfam; PF00313; CSD; 1.
DR   Pfam; PF00098; zf-CCHC; 2.
DR   PRINTS; PR00939; C2HCZNFINGER.
DR   PRINTS; PR00050; COLDSHOCK.
DR   ProDom; PD000621; Cold_shock; 1.
DR   SMART; SM00357; CSP; 1.
DR   SMART; SM00343; ZnF_C2HC; 2.
DR   PROSITE; PS00352; COLD_SHOCK; 1.
DR   PROSITE; PS50158; ZF_CCHC; 2.
KW   RNA-binding.
SQ   SEQUENCE   205 AA;  19225 MW;  0918778B79CD0B51 CRC64;

  Query Match             31.7%;  Score 97;  DB 2;  Length 205;
  Best Local Similarity   51.2%;  Pred. No. 0.035;
  Matches   21;  Conservative    2;  Mismatches    6;  Indels   12;  Gaps    2;

Qy          1 GGGGGGGTYSCHFGPLTWVCKPQGGGGGGGGTYSC----HF 37
              |||||||  |    |         :|||||||| :||    ||
Db        164 GGGGGGGGYGGGGG--------RGGGGGGGGCFSCGESGHF 196


RESULT 4
Q8LPA7_WHEAT
ID   Q8LPA7_WHEAT PRELIMINARY;      PRT;   229 AA.
AC   Q8LPA7;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Cold shock protein-1.
GN   Name=WCSP1;
OS   Triticum aestivum (Wheat).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC   Triticeae; Triticum.
OX   NCBI_TaxID=4565;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22218063; PubMed=12122010; DOI=10.1074/jbc.M205774200;
RA   Karlson D., Nakaminami K., Toyomasu T., Imai R.;
RT   "A cold-regulated nucleic acid-binding protein of winter wheat shares
RT   a domain with bacterial cold shock proteins.";
RL   J. Biol. Chem. 277:35248-35256(2002).
CC   -!- SIMILARITY: Contains 1 CSD (cold-shock) domain.
DR   EMBL; AB066265; BAB78536.2; -; mRNA.
DR   HSSP; P15277; 1MJC.
DR   Gramene; Q8LPA7; -.
DR   GO; GO:0003677; F:DNA binding; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR011129; CSP.
DR   InterPro; IPR002059; CSP_DNA_bd.
DR   InterPro; IPR012340; OB_NA_bd_sub.
DR   InterPro; IPR001878; Znf_CCHC.
DR   Pfam; PF00313; CSD; 1.
DR   Pfam; PF00098; zf-CCHC; 3.
DR   PRINTS; PR00939; C2HCZNFINGER.
DR   PRINTS; PR00050; COLDSHOCK.
DR   ProDom; PD000621; Cold_shock; 1.
DR   SMART; SM00357; CSP; 1.
DR   SMART; SM00343; ZnF_C2HC; 3.
DR   PROSITE; PS00352; COLD_SHOCK; 1.
DR   PROSITE; PS50158; ZF_CCHC; 3.
KW   RNA-binding.
SQ   SEQUENCE   229 AA;  21384 MW;  4CB5C9B6323BD23C CRC64;

  Query Match             31.7%;  Score 97;  DB 2;  Length 229;
  Best Local Similarity   51.2%;  Pred. No. 0.039;
  Matches   21;  Conservative    2;  Mismatches    6;  Indels   12;  Gaps    2;

Qy          1 GGGGGGGTYSCHFGPLTWVCKPQGGGGGGGGTYSC----HF 37
              |||||||  |    |         :|||||||| :||    ||
Db        188 GGGGGGGGYGGGGG--------RGGGGGGGGCFSCGESGHF 220


RESULT 5
Q6YUR8_ORYSA
ID   Q6YUR8_ORYSA PRELIMINARY;      PRT;   241 AA.
AC   Q6YUR8;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE   Putative Glycine-rich protein 2.
GN   Name=OSJNBb0088N06.21; Synonyms=OJ1020_C02.12;
OS   Oryza sativa (japonica cultivar-group).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=39947;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Sasaki T., Matsumoto T., Katayose Y.;
RT   "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 2, BAC
RT   clone:OSJNBb0088N06.";
RL   Submitted (OCT-2002) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RA   Sasaki T., Matsumoto T., Yamamoto K.;
RT   "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 2, BAC
RT   clone:OJ1020_C02.";
RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Contains 1 CSD (cold-shock) domain.
DR   EMBL; AP005851; BAD08139.1; -; Genomic_DNA.
DR   EMBL; AP004078; BAD07599.1; -; Genomic_DNA.
DR   Gramene; Q6YUR8; -.
```

```
DR   GO; GO:0003677; F:DNA binding; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR011129; CSP.
DR   InterPro; IPR002059; CSP_DNA_bd.
DR   InterPro; IPR002952; Eggshell.
DR   InterPro; IPR012340; OB_NA_bd_sub.
DR   InterPro; IPR001878; Znf_CCHC.
DR   Pfam; PF00313; CSD; 1.
DR   Pfam; PF00098; zf-CCHC; 4.
DR   PRINTS; PR00939; C2HCZNFINGER.
DR   PRINTS; PR00050; COLDSHOCK.
DR   PRINTS; PR01228; EGGSHELL.
DR   ProDom; PD000621; Cold_shock; 1.
DR   SMART; SM00357; CSP; 1.
DR   SMART; SM00343; ZnF_C2HC; 4.
DR   PROSITE; PS00352; COLD_SHOCK; 1.
DR   PROSITE; PS50158; ZF_CCHC; 4.
KW   RNA-binding.
SQ   SEQUENCE   241 AA;  22723 MW;  69E6A187A7B35E03 CRC64;

  Query Match             31.2%;  Score 95.5;  DB 2;  Length 241;
  Best Local Similarity   41.4%;  Pred. No. 0.058;
  Matches   24;  Conservative    3;  Mismatches   22;  Indels    9;  Gaps    3;

Qy          1 GGGGGGGTYSC-HFGPLTWVCKPQGGGGGGG-------TYSC-HFGPLTWVCKPQGG 49
              ||| ||| : |   | :   |   ||||||||        |:|   | |   |   ||
Db        154 GGGAGGGCFKCGEMGHMARDCFNSGGGGGGGGGGGGGGGACYNCGETGHLARDCYNGGG 211


RESULT 6
Q75QN8_WHEAT
ID   Q75QN8_WHEAT PRELIMINARY;      PRT;   231 AA.
AC   Q75QN8;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Cold shock domain protein 3.
GN   Name=WCSP3;
OS   Triticum aestivum (Wheat).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC   Triticeae; Triticum.
OX   NCBI_TaxID=4565;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Nakaminami K., Imai R.;
RL   Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Contains 1 CSD (cold-shock) domain.
DR   EMBL; AB161683; BAD08701.1; -; mRNA.
DR   Gramene; Q75QN8; -.
DR   GO; GO:0003677; F:DNA binding; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR011129; CSP.
DR   InterPro; IPR002059; CSP_DNA_bd.
DR   InterPro; IPR012340; OB_NA_bd_sub.
DR   InterPro; IPR001878; Znf_CCHC.
DR   Pfam; PF00313; CSD; 1.
DR   Pfam; PF00098; zf-CCHC; 3.
DR   PRINTS; PR00939; C2HCZNFINGER.
DR   PRINTS; PR00050; COLDSHOCK.
DR   ProDom; PD000621; Cold_shock; 1.
DR   SMART; SM00357; CSP; 1.
DR   SMART; SM00343; ZnF_C2HC; 3.
DR   PROSITE; PS00352; COLD_SHOCK; 1.
DR   PROSITE; PS50158; ZF_CCHC; 3.
KW   RNA-binding.
SQ   SEQUENCE   231 AA;  21544 MW;  F6E1FE0104CDE2C6 CRC64;

  Query Match             31.0%;  Score 95;  DB 2;  Length 231;
  Best Local Similarity   41.5%;  Pred. No. 0.062;
  Matches   27;  Conservative    4;  Mismatches   16;  Indels   18;  Gaps    5;

Qy          1 GGGGGGGTYSC-HFGPLTWVCKPQGGGGGGG--------------GTYSC-HFGPLTWVC 44
              ||||| | | |   | ::  | |||||||||                  | |   | ::  |
Db        131 GGGGGRGCYKCGEDGHISRDC-PQGGGGGGGYGGGGYGGGGGGGRECYKCGEEGHISRDC 189

Qy         45 KPQGG 49
              ||||
Db        190 -PQGG 193


RESULT 7
Q84WQ1_ARATH
ID   Q84WQ1_ARATH PRELIMINARY;      PRT;   473 AA.
AC   Q84WQ1;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Hypothetical protein At5g19090 (Fragment).
GN   Name=At5g19090;
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Yamada K., Chan M.M., Chang C.H., Dale J.M., Hsuan V.W., Lee J.M.,
RA   Onodera C.S., Quach H.L., Tang C., Toriumi M., Wong C., Wu H.C.,
RA   Yu G., Yuan S., Carninci P., Chen H., Cheuk R., Hayashizaki Y.,
RA   Ishida J., Jones T., Kamiya A., Kawai J., Kim C.J., Narusaka M.,
RA   Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M., Shinn P.,
RA   Southwick A., Tripp M.G., Wu T., Shinozaki K., Davis R.W., Ecker J.R.,
RA   Theologis A.;
RL   Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BT002910; AAO22726.1; -; mRNA.
KW   Hypothetical protein.
FT   NON_TER       1      1
SQ   SEQUENCE   473 AA;  47613 MW;  EB10A0539821C568 CRC64;

  Query Match             30.7%;  Score 94;  DB 2;  Length 473;
  Best Local Similarity   49.0%;  Pred. No. 0.16;
  Matches   24;  Conservative    3;  Mismatches   12;  Indels   10;  Gaps    3;

Qy          1 GGGGGGGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGG 49
              |||||||  |   | |    :| ||||||||      || : :  | ||
Db        276 GGGGGGGPMS---GGLPPGFRPMGGGGGGGG------GPQS-MSMPMGG 314


RESULT 8
Q7XGH3_ORYSA
ID   Q7XGH3_ORYSA PRELIMINARY;      PRT;   249 AA.
AC   Q7XGH3;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein.
GN   ORFNames=OSJNBb0081F12.20;
OS   Oryza sativa (japonica cultivar-group).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=39947;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   The Rice Chromosome 10 Sequencing Consortium;
RT   "In-depth view of structure, activity, and evolution of rice
RT   chromosome 10.";
RL   Science 300:1566-1569(2003).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RA   Buell C.R., Wing R.A., McCombie W.R., Messing J., Yuan Q.;
RL   Submitted (MAY-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE017062; AAP52328.1; -; Genomic_DNA.
DR   Gramene; Q7XGH3; -.
KW   Hypothetical protein.
```

```
SQ   SEQUENCE   249 AA;  25708 MW;  987D060FF7EDD420 CRC64;

  Query Match             30.4%;  Score 93;  DB 2;  Length 249;
  Best Local Similarity   40.0%;  Pred. No. 0.11;
  Matches   22;  Conservative    6;  Mismatches   21;  Indels    6;  Gaps    2;

Qy          1 GGGGGGGTYSCHFGPL-TWVCKPQGGGGGGGGTYSCHF-----GPLTWVCKPQGG 49
              |||||||  | :   :  ||   :| ||||||  : |      | :   |  :||
Db         49 GGGGGGGADSWNLEEIGLWVRSAEGNGGGGGGEDNGHCVDDEGGAVGAACPGEGG 103


RESULT 9
Q8S748_ORYSA
ID   Q8S748_ORYSA PRELIMINARY;      PRT;   249 AA.
AC   Q8S748;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein OSJNBb0081F12.20.
GN   Name=OSJNBb0081F12.20;
OS   Oryza sativa (Rice).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=4530;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Nipponbare;
RA   McCombie W.R., de la Bastide M., Spiegel L., Kirchoff K., Preston R.,
RA   Kuit K., Nascimento L., Bell M., Balija V., Baker J., Vil M.D.,
RA   Zutavern T., Santos L., Miller B., Cunnius D.M., Shah R., King L.,
RA   Bahret A., Yang C., Dike S., O'Shaughnessy A., Palmer L., Dedhia N.;
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC090488; AAM01020.1; -; Genomic_DNA.
DR   Gramene; Q8S748; -.
KW   Hypothetical protein.
SQ   SEQUENCE   249 AA;  25708 MW;  987D060FF7EDD420 CRC64;

  Query Match             30.4%;  Score 93;  DB 2;  Length 249;
  Best Local Similarity   40.0%;  Pred. No. 0.11;
  Matches   22;  Conservative    6;  Mismatches   21;  Indels    6;  Gaps    2;

Qy          1 GGGGGGGTYSCHFGPL-TWVCKPQGGGGGGGGTYSCHF-----GPLTWVCKPQGG 49
              |||||||  | :   :  ||   :| ||||||  : |      | :   |  :||
Db         49 GGGGGGGADSWNLEEIGLWVRSAEGNGGGGGGEDNGHCVDDEGGAVGAACPGEGG 103


RESULT 10
Q61PL7_CAEBR
ID   Q61PL7_CAEBR PRELIMINARY;      PRT;   318 AA.
AC   Q61PL7;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Hypothetical protein CBG07448.
GN   Name=CBG07448;
OS   Caenorhabditis briggsae.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6238;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RG   The C.briggsae Sequencing Consortium;
RL   Submitted (SEP-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; CAAC01000032; CAE63146.1; -; Genomic_DNA.
DR   GO; GO:0005737; C:cytoplasm; IEA.
DR   GO; GO:0042302; F:structural constituent of cuticle; IEA.
DR   GO; GO:0006817; P:phosphate transport; IEA.
DR   InterPro; IPR008160; Collagen.
DR   InterPro; IPR002486; Col_cuticle_N.
DR   Pfam; PF01391; Collagen; 2.
DR   Pfam; PF01484; Col_cuticle_N; 1.
KW   Collagen; Hypothetical protein.

SQ   SEQUENCE   318 AA;  29628 MW;  82F745AC2C320B2F CRC64;

  Query Match             30.2%;  Score 92.5;  DB 2;  Length 318;
  Best Local Similarity   47.8%;  Pred. No. 0.15;
  Matches   22;  Conservative    0;  Mismatches   13;  Indels   11;  Gaps    2;

Qy          1 GGGGGGGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKP 46
              || |||| ||   |      |||||||||   |     | | |
Db         85 GGAGGGGGYSAGGG------GGQGGGGGGGGDGGC-----TGCCNP 119


RESULT 11
Q84UR8_ORYSA
ID   Q84UR8_ORYSA PRELIMINARY;      PRT;   197 AA.
AC   Q84UR8;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Putative cold shock protein-1.
GN   Name=P0582D05.112;
OS   Oryza sativa (japonica cultivar-group).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=39947;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Sasaki T., Matsumoto T., Yamamoto K.;
RL   Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Contains 1 CSD (cold-shock) domain.
DR   EMBL; AP004591; BAC66711.1; -; Genomic_DNA.
DR   HSSP; P15277; 1MJC.
DR   Gramene; Q84UR8; -.
DR   GO; GO:0003677; F:DNA binding; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR011129; CSP.
DR   InterPro; IPR002059; CSP_DNA_bd.
DR   InterPro; IPR012340; OB_NA_bd_sub.
DR   InterPro; IPR001878; ZnF_CCHC.
DR   Pfam; PF00313; CSD; 1.
DR   Pfam; PF00098; zf-CCHC; 2.
DR   PRINTS; PR00939; C2HCZNFINGER.
DR   PRINTS; PR00050; COLDSHOCK.
DR   ProDom; PD000621; Cold_shock; 1.
DR   SMART; SM00357; CSP; 1.
DR   SMART; SM00343; ZnF_C2HC; 2.
DR   PROSITE; PS00352; COLD_SHOCK; 1.
DR   PROSITE; PS50158; ZF_CCHC; 2.
KW   RNA-binding.
SQ   SEQUENCE   197 AA;  18694 MW;  6DD7597CDCE0C70B CRC64;

  Query Match             29.6%;  Score 90.5;  DB 2;  Length 197;
  Best Local Similarity   35.7%;  Pred. No. 0.15;
  Matches   20;  Conservative    6;  Mismatches    9;  Indels   21;  Gaps    2;

Qy          1 GGGGGGGTYSCH----FGPLTWVC----------------KPQGGGGGGGGTYSC 35
              |||||||: :|:     | :  |                 : |||||||| |:|
Db        127 GGGGGGGSRACYKCGEEGHMARDCSQGGGGGGGYGGGGGGYRGGGGGGGGGGCYNC 182


RESULT 12
Q8L7T1_ARATH
ID   Q8L7T1_ARATH PRELIMINARY;      PRT;   225 AA.
AC   Q8L7T1;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   At2g41260/F13H10.19.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
```

```
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Shinn P., Chen H., Cheuk R., Kim C.J., Banh J., Bowser L.,
RA   Carninci P., Chang E., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA   Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,
RA   Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,
RA   Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA   Tang C.C., Toriumi M., Wu H.C., Yamada K., Yamamura Y., Yu G., Yu S.,
RA   Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RL   Submitted (JUN-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY128274; AAM91083.1; -; mRNA.
SQ   SEQUENCE   225 AA;  23887 MW;  B2CFF845C3859576 CRC64;

  Query Match             29.6%;  Score 90.5;  DB 2;  Length 225;
  Best Local Similarity   30.4%;  Pred. No. 0.17;
  Matches   21;  Conservative    4;  Mismatches   13;  Indels   31;  Gaps    2;

Qy          1 GGGGGGGTYSCHF--------GPLTWVCKPQ----------------------GGGGGG 29
              |||||||   | :        |   : |: |                      ||||||
Db        123 GGGGGGGRGGCRWGCCGGWWRGRCRYCCRSQAEASEVVETVEPNDVEPQQGGRGGGGGGG 182

Qy         30 GGTYSCHFG 38
              ||   | :|
Db        183 GGRGGCRWG 191


RESULT 13
Q9S7S3_ARATH
ID   Q9S7S3_ARATH PRELIMINARY;      PRT;   280 AA.
AC   Q9S7S3;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE   Late embryogenesis abundant M17 protein.
GN   Name=At2g41260; Synonyms=M17;
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Rounsley S.D., Lin X., Kaul S., Shea T.P., Fujii C.Y., Mason T.M.,
RA   Shen M., Ronning C.M., Fraser C.M., Somerville C.R., Venter J.C.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RA   Town C.D., Kaul S.;
RL   Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=99320877; PubMed=10394954; DOI=10.1023/A:1026403215270;
RA   Raynal M., Guilleminot J., Gueguen C., Cooke R., Delseny M.,
RA   Gruber V.;
RT   "Structure, organization and expression of two closely related novel
RT   Lea (late-embryogenesis-abundant) genes in Arabidopsis thaliana.";
RL   Plant Mol. Biol. 40:153-165(1999).
DR   EMBL; AC005662; AAC78545.1; -; Genomic_DNA.
DR   EMBL; AF076979; AAC27641.1; -; Genomic_DNA.
DR   PIR; G84839; G84839.
SQ   SEQUENCE   280 AA;  29559 MW;  52E91864C8035BF6 CRC64;

  Query Match             29.6%;  Score 90.5;  DB 2;  Length 280;
  Best Local Similarity   30.4%;  Pred. No. 0.21;
  Matches   21;  Conservative    4;  Mismatches   13;  Indels   31;  Gaps    2;

Qy          1 GGGGGGGTYSCHF--------GPLTWVCKPQ----------------------GGGGGG 29
              |||||||   | :        |   : |: |                      ||||||
Db        123 GGGGGGGRGGCRWGCCGGWWRGRCRYCCRSQAEASEVVETVEPNDVEPQQGGRGGGGGGG 182

Qy         30 GGTYSCHFG 38
              ||   | :|
Db        183 GGRGGCRWG 191


RESULT 14
Q9VD49_DROME
ID   Q9VD49_DROME PRELIMINARY;      PRT;   117 AA.
AC   Q9VD49;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   CG5778-PA (GH13168p).
GN   ORFNames=CG5778, CG5778;
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22426065; PubMed=12537568;
RA   Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA   Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA   George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,
RA   Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA   Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA   Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT   "Finishing a whole-genome shotgun: release 3 of the Drosophila
RT   melanogaster euchromatic genome sequence.";
RL   Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22426070; PubMed=12537573;
RA   Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J.W., Svirskas R.,
```

```
RA   Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA   Ashburner M., Celniker S.E.;
RT   "The transposable elements of the Drosophila melanogaster euchromatin:
RT   a genomics perspective.";
RL   Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).
RN   [4]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22426069; PubMed=12537572;
RA   Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA   Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA   Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA   Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA   Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA   Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA   Lewis S.E.;
RT   "Annotation of the Drosophila melanogaster euchromatic genome: a
RT   systematic review.";
RL   Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN   [5]
RP   NUCLEOTIDE SEQUENCE.
RG   Berkeley Drosophila Genome Project;
RA   Celniker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,
RA   Hoskins R., Stapleton M., Pacleb J., Park S., Svirskas R., Smith E.,
RA   Yu C., Rubin G.;
RT   "Drosophila melanogaster release 4 sequence.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   NUCLEOTIDE SEQUENCE.
RG   FlyBase;
RL   Submitted (MAR-2005) to the EMBL/GenBank/DDBJ databases.
RN   [7]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Berkeley;
RA   Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA   Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA   George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA   Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA   Patel S., Phouanenavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA   Celniker S.;
RL   Submitted (JUN-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE003737; AAF55954.1; -; Genomic_DNA.
DR   EMBL; AY118781; AAM50641.1; -; mRNA.
DR   Ensembl; CG5778; Drosophila melanogaster.
DR   FlyBase; FBgn0038930; CG5778.
SQ   SEQUENCE   117 AA;  10632 MW;  3AE2D79CA8924A96 CRC64;

  Query Match             29.4%;  Score 90;  DB 2;  Length 117;
  Best Local Similarity   48.8%;  Pred. No. 0.1;
  Matches   20;  Conservative    2;  Mismatches    3;  Indels   16;  Gaps    2;

Qy          2 GGGGGGTYSCHFGPLTWVCKPQGGGGGGGG---TYSCHFGP 39
              |||||| |:             ||||||||    || :|||
Db         58 GGGGGGYYN------------GGGGGGGGRRPVYSGNFGP 85


RESULT 15
GRP2_NICSY
ID   GRP2_NICSY     STANDARD;      PRT;   214 AA.
AC   P27484;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   10-MAY-2005 (Rel. 47, Last annotation update)
DE   Glycine-rich protein 2.
GN   Name=GRP-2;
OS   Nicotiana sylvestris (Wood tobacco).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC   asterids; lamiids; Solanales; Solanaceae; Nicotiana.
OX   NCBI_TaxID=4096;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=92003709; PubMed=1912512;
RA   Obokata J., Ohme M., Hayashida N.;
RT   "Nucleotide sequence of a cDNA clone encoding a putative glycine-rich
RT   protein of 19.7 kDa in Nicotiana sylvestris.";
RL   Plant Mol. Biol. 17:953-955(1991).
CC   -!- SIMILARITY: Contains 2 C2HC-type zinc fingers.
CC   -!- SIMILARITY: Contains 2 CCHC-type zinc fingers.
CC   -!- SIMILARITY: Contains 1 CSD (cold-shock) domain.
CC   -!- CAUTION: Was originally (Ref.1) thought to be a cell wall
CC       structural protein.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; X60007; CAA42622.1; -; mRNA.
DR   PIR; S17731; KNNT2S.
DR   HSSP; O54310; 1G6P.
DR   InterPro; IPR011129; CSP.
DR   InterPro; IPR002059; CSP_DNA_bd.
DR   InterPro; IPR012340; OB_NA_bd_sub.
DR   InterPro; IPR001878; Znf_CCHC.
DR   Pfam; PF00313; CSD; 1.
DR   Pfam; PF00098; zf-CCHC; 2.
DR   PRINTS; PR00939; C2HCZNFINGER.
DR   PRINTS; PR00050; COLDSHOCK.
DR   ProDom; PD000621; Cold_shock; 1.
DR   SMART; SM00357; CSP; 1.
DR   SMART; SM00343; ZnF_C2HC; 2.
DR   PROSITE; PS00352; COLD_SHOCK; 1.
DR   PROSITE; PS50158; ZF_CCHC; 2.
KW   Metal-binding; Repeat; RNA-binding; Zinc; Zinc-finger.
FT   DOMAIN        8     75       CSD.
FT   ZN_FING     157    174       CCHC-type 1.
FT   ZN_FING     194    211       CCHC-type 2.
FT   COMPBIAS     82    158       Gly-rich.
FT   COMPBIAS    176    195       Gly-rich.
SQ   SEQUENCE   214 AA;  19746 MW;  E28DB84538F2A0AA CRC64;

  Query Match             29.2%;  Score 89.5;  DB 1;  Length 214;
  Best Local Similarity   51.2%;  Pred. No. 0.2;
  Matches   21;  Conservative    0;  Mismatches    5;  Indels   15;  Gaps    3;

Qy          1 GGGGGGGTYSCHFGPLTWVCKPQGGGGGGGGTYSC----HF 37
              |||||||     ||          |||||||| | |    ||
Db        176 GGGGGGG----RFG-------GGGGGGGGGGCYKCGEDGHF 205


Search completed: March 31, 2006, 16:35:16
Job time : 185.567 secs
```

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: March 31, 2006, 16:35:37 ; Search time 48.0249 Seconds
(without alignments)
84.354 Million cell updates/sec

Title: US-10-609-217-339
Perfect score: 306
Sequence: 1 GGGGGGGTYSCHFGPLTWVC.....GGTYSCHFGPLTWVCKPQGG 49

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/6_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/H_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/RE_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	306	100.0	49	2	US-09-428-082B-339	Sequence 339, App
2	306	100.0	277	2	US-09-428-082B-22	Sequence 22, Appl
3	276	90.2	49	2	US-09-428-082B-340	Sequence 340, App
4	276	90.2	57	2	US-09-428-082B-417	Sequence 417, App
5	276	90.2	277	2	US-09-428-082B-20	Sequence 20, Appl
6	240	78.4	40	2	US-09-428-082B-92	Sequence 92, Appl
7	239.5	78.3	46	2	US-09-428-082B-95	Sequence 95, Appl
8	231	75.5	39	2	US-09-428-082B-395	Sequence 395, App
9	192	62.7	36	2	US-09-428-082B-403	Sequence 403, App
10	168.5	55.1	253	2	US-09-428-082B-16	Sequence 16, Appl
11	161.5	52.8	253	2	US-09-428-082B-18	Sequence 18, Appl
12	156	51.0	25	2	US-09-428-082B-1034	Sequence 1034, Ap
13	126	41.2	20	1	US-08-484-135-8	Sequence 8, Appli
14	126	41.2	20	1	US-08-484-135-20	Sequence 20, Appl
15	126	41.2	20	1	US-08-484-135-42	Sequence 42, Appl
16	126	41.2	20	1	US-08-484-635-8	Sequence 8, Appli
17	126	41.2	20	1	US-08-484-631-8	Sequence 8, Appli
18	126	41.2	20	1	US-08-641-071-1	Sequence 1, Appli
19	126	41.2	20	1	US-08-827-570-8	Sequence 8, Appli
20	126	41.2	20	2	US-08-905-310-2	Sequence 2, Appli
21	126	41.2	20	2	US-08-825-852-34	Sequence 34, Appl
22	126	41.2	20	2	US-08-786-690-1	Sequence 1, Appli
23	126	41.2	20	2	US-08-786-690-4	Sequence 4, Appli
24	126	41.2	20	2	US-09-052-888-34	Sequence 34, Appl
25	126	41.2	20	2	US-09-723-890-34	Sequence 34, Appl
26	126	41.2	20	2	US-09-723-901-34	Sequence 34, Appl
27	126	41.2	20	2	US-09-723-547-34	Sequence 34, Appl
28	126	41.2	20	2	US-09-724-127-34	Sequence 34, Appl
29	126	41.2	20	2	US-09-723-931-34	Sequence 34, Appl
30	126	41.2	20	2	US-09-428-082B-87	Sequence 87, Appl
31	126	41.2	20	2	US-09-428-082B-93	Sequence 93, Appl
32	126	41.2	20	2	US-09-428-082B-1025	Sequence 1025, Ap
33	126	41.2	20	2	US-09-723-873-34	Sequence 34, Appl
34	126	41.2	20	2	US-09-724-114-34	Sequence 34, Appl
35	126	41.2	20	2	US-09-723-913-34	Sequence 34, Appl
36	126	41.2	20	2	US-09-723-912-34	Sequence 34, Appl
37	126	41.2	20	2	US-09-724-095-34	Sequence 34, Appl
38	126	41.2	20	2	US-09-724-157-34	Sequence 34, Appl
39	126	41.2	20	2	US-09-724-062-34	Sequence 34, Appl
40	126	41.2	20	2	US-09-724-065-34	Sequence 34, Appl
41	126	41.2	20	2	US-09-724-481-34	Sequence 34, Appl
42	126	41.2	22	2	US-09-428-082B-97	Sequence 97, Appl
43	126	41.2	23	1	US-08-484-635-20	Sequence 20, Appl
44	126	41.2	23	1	US-08-484-631-20	Sequence 20, Appl
45	126	41.2	23	1	US-08-827-570-20	Sequence 20, Appl

ALIGNMENTS

RESULT 1
US-09-428-082B-339
; Sequence 339, Application US/09428082B
; Patent No. 6660843
; GENERAL INFORMATION:
; APPLICANT: FEIGE, ULRICH
; APPLICANT: LIU, CHUAN-FA
; APPLICANT: CHEETHAM, JANET C.
; APPLICANT: BOONE, THOMAS CHARLES
; TITLE OF INVENTION: MODIFIED PEPTIDES AS THERAPEUTIC AGENTS
; FILE REFERENCE: A-527
; CURRENT APPLICATION NUMBER: US/09/428,082B
; CURRENT FILING DATE: 1999-10-22
; PRIOR APPLICATION NUMBER: 60/105,371
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 1133
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 339
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: EPO-MIMETIC
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: Fc domain attached at Position 1 of the N-terminus
US-09-428-082B-339

Query Match 100.0%; Score 306; DB 2; Length 49;
Best Local Similarity 100.0%; Pred. No. 3.3e-25;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GGGGGGGTYSCHFGPLTWVCKPQGGGGGGGGGTYSCHFGPLTWVCKPQGG 49
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGGGGGTYSCHFGPLTWVCKPQGGGGGGGGGTYSCHFGPLTWVCKPQGG 49
```

RESULT 2
US-09-428-082B-22
; Sequence 22, Application US/09428082B
; Patent No. 6660843
; GENERAL INFORMATION:
; APPLICANT: FEIGE, ULRICH
; APPLICANT: LIU, CHUAN-FA
; APPLICANT: CHEETHAM, JANET C.
; APPLICANT: BOONE, THOMAS CHARLES
; TITLE OF INVENTION: MODIFIED PEPTIDES AS THERAPEUTIC AGENTS
; FILE REFERENCE: A-527
; CURRENT APPLICATION NUMBER: US/09/428,082B

```
;  CURRENT FILING DATE:  1999-10-22
;  PRIOR APPLICATION NUMBER: 60/105,371
;  PRIOR FILING DATE: 1998-10-23
;  NUMBER OF SEQ ID NOS: 1133
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22
;   LENGTH: 277
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Fc-EMP-EMP
US-09-428-082B-22

  Query Match             100.0%;  Score 306;  DB 2;  Length 277;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-24;
  Matches   49;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGGGGGGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGG 49
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db        229 GGGGGGGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGG 277


RESULT 3
US-09-428-082B-340
; Sequence 340, Application US/09428082B
; Patent No. 6660843
; GENERAL INFORMATION:
;  APPLICANT: FEIGE, ULRICH
;  APPLICANT:  LIU, CHUAN-FA
;  APPLICANT:  CHEETHAM, JANET C.
;  APPLICANT:  BOONE, THOMAS CHARLES
;  TITLE OF INVENTION: MODIFIED PEPTIDES AS THERAPEUTIC AGENTS
;  FILE REFERENCE: A-527
;  CURRENT APPLICATION NUMBER: US/09/428,082B
;  CURRENT FILING DATE:  1999-10-22
;  PRIOR APPLICATION NUMBER: 60/105,371
;  PRIOR FILING DATE: 1998-10-23
;  NUMBER OF SEQ ID NOS: 1133
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 340
;   LENGTH: 49
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: EPO-MIMETIC
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (49)..(49)
;   OTHER INFORMATION: Fc domain attached at Position 49 of the C-terminus
US-09-428-082B-340

  Query Match             90.2%;  Score 276;  DB 2;  Length 49;
  Best Local Similarity   100.0%;  Pred. No. 4.2e-22;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 GGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGG 49
              ||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGG 44


RESULT 4
US-09-428-082B-417
; Sequence 417, Application US/09428082B
; Patent No. 6660843
; GENERAL INFORMATION:
;  APPLICANT: FEIGE, ULRICH
;  APPLICANT:  LIU, CHUAN-FA
;  APPLICANT:  CHEETHAM, JANET C.
;  APPLICANT:  BOONE, THOMAS CHARLES
;  TITLE OF INVENTION: MODIFIED PEPTIDES AS THERAPEUTIC AGENTS
;  FILE REFERENCE: A-527
;  CURRENT APPLICATION NUMBER: US/09/428,082B
;  CURRENT FILING DATE:  1999-10-22
;  PRIOR APPLICATION NUMBER: 60/105,371
;  PRIOR FILING DATE: 1998-10-23
;  NUMBER OF SEQ ID NOS: 1133
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 417
;   LENGTH: 57
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: EMP-EMP-Fc
US-09-428-082B-417

  Query Match             90.2%;  Score 276;  DB 2;  Length 57;
  Best Local Similarity   100.0%;  Pred. No. 4.9e-22;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 GGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGG 49
              ||||||||||||||||||||||||||||||||||||||||||||
Db          2 GGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGG 45


RESULT 5
US-09-428-082B-20
; Sequence 20, Application US/09428082B
; Patent No. 6660843
; GENERAL INFORMATION:
;  APPLICANT: FEIGE, ULRICH
;  APPLICANT:  LIU, CHUAN-FA
;  APPLICANT:  CHEETHAM, JANET C.
;  APPLICANT:  BOONE, THOMAS CHARLES
;  TITLE OF INVENTION: MODIFIED PEPTIDES AS THERAPEUTIC AGENTS
;  FILE REFERENCE: A-527
;  CURRENT APPLICATION NUMBER: US/09/428,082B
;  CURRENT FILING DATE:  1999-10-22
;  PRIOR APPLICATION NUMBER: 60/105,371
;  PRIOR FILING DATE: 1998-10-23
;  NUMBER OF SEQ ID NOS: 1133
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
;   LENGTH: 277
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: EMP-EMP-Fc
US-09-428-082B-20

  Query Match             90.2%;  Score 276;  DB 2;  Length 277;
  Best Local Similarity   100.0%;  Pred. No. 2.3e-21;
  Matches   44;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 GGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGG 49
              ||||||||||||||||||||||||||||||||||||||||||||
Db          2 GGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGG 45


RESULT 6
US-09-428-082B-92
; Sequence 92, Application US/09428082B
; Patent No. 6660843
; GENERAL INFORMATION:
;  APPLICANT: FEIGE, ULRICH
;  APPLICANT:  LIU, CHUAN-FA
;  APPLICANT:  CHEETHAM, JANET C.
;  APPLICANT:  BOONE, THOMAS CHARLES
;  TITLE OF INVENTION: MODIFIED PEPTIDES AS THERAPEUTIC AGENTS
;  FILE REFERENCE: A-527
;  CURRENT APPLICATION NUMBER: US/09/428,082B
;  CURRENT FILING DATE:  1999-10-22
;  PRIOR APPLICATION NUMBER: 60/105,371
;  PRIOR FILING DATE: 1998-10-23
;  NUMBER OF SEQ ID NOS: 1133
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 92
```

```
;   LENGTH: 40
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: EPO-MIMETIC PEPTIDE
US-09-428-082B-92

  Query Match             78.4%;  Score 240;  DB 2;  Length 40;
  Best Local Similarity   90.9%;  Pred. No. 1.8e-18;
  Matches   40;  Conservative    0;  Mismatches    0;  Indels    4;  Gaps    1;

Qy          6 GGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGG 49
              |||||||||||||||||    ||||||||||||||||||||||
Db          1 GGTYSCHFGPLTWVCKPQ----GGGGTYSCHFGPLTWVCKPQGG 40


RESULT 7
US-09-428-082B-95
; Sequence 95, Application US/09428082B
; Patent No. 6660843
; GENERAL INFORMATION:
;  APPLICANT: FEIGE, ULRICH
;  APPLICANT:  LIU, CHUAN-FA
;  APPLICANT:  CHEETHAM, JANET C.
;  APPLICANT:  BOONE, THOMAS CHARLES
;  TITLE OF INVENTION: MODIFIED PEPTIDES AS THERAPEUTIC AGENTS
;  FILE REFERENCE: A-527
;  CURRENT APPLICATION NUMBER: US/09/428,082B
;  CURRENT FILING DATE:  1999-10-22
;  PRIOR APPLICATION NUMBER: 60/105,371
;  PRIOR FILING DATE: 1998-10-23
;  NUMBER OF SEQ ID NOS: 1133
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 95
;   LENGTH: 46
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: EPO-MIMETIC PEPTIDE
US-09-428-082B-95

  Query Match             78.3%;  Score 239.5;  DB 2;  Length 46;
  Best Local Similarity   90.9%;  Pred. No. 2.4e-18;
  Matches   40;  Conservative    0;  Mismatches    3;  Indels    1;  Gaps    1;

Qy          6 GGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGG 49
              |||||||||||||||||  ||   ||||||||||||||||||||
Db          1 GGTYSCHFGPLTWVCKPQ-GGSSKGGTYSCHFGPLTWVCKPQGG 43


RESULT 8
US-09-428-082B-395
; Sequence 395, Application US/09428082B
; Patent No. 6660843
; GENERAL INFORMATION:
;  APPLICANT: FEIGE, ULRICH
;  APPLICANT:  LIU, CHUAN-FA
;  APPLICANT:  CHEETHAM, JANET C.
;  APPLICANT:  BOONE, THOMAS CHARLES
;  TITLE OF INVENTION: MODIFIED PEPTIDES AS THERAPEUTIC AGENTS
;  FILE REFERENCE: A-527
;  CURRENT APPLICATION NUMBER: US/09/428,082B
;  CURRENT FILING DATE:  1999-10-22
;  PRIOR APPLICATION NUMBER: 60/105,371
;  PRIOR FILING DATE: 1998-10-23
;  NUMBER OF SEQ ID NOS: 1133
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 395
;   LENGTH: 39
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Fc-EMP
US-09-428-082B-395

  Query Match             75.5%;  Score 231;  DB 2;  Length 39;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-17;
  Matches   37;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGGGGGGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHF 37
              |||||||||||||||||||||||||||||||||||||||
Db          3 GGGGGGGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHF 39


RESULT 9
US-09-428-082B-403
; Sequence 403, Application US/09428082B
; Patent No. 6660843
; GENERAL INFORMATION:
;  APPLICANT: FEIGE, ULRICH
;  APPLICANT:  LIU, CHUAN-FA
;  APPLICANT:  CHEETHAM, JANET C.
;  APPLICANT:  BOONE, THOMAS CHARLES
;  TITLE OF INVENTION: MODIFIED PEPTIDES AS THERAPEUTIC AGENTS
;  FILE REFERENCE: A-527
;  CURRENT APPLICATION NUMBER: US/09/428,082B
;  CURRENT FILING DATE:  1999-10-22
;  PRIOR APPLICATION NUMBER: 60/105,371
;  PRIOR FILING DATE: 1998-10-23
;  NUMBER OF SEQ ID NOS: 1133
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 403
;   LENGTH: 36
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: EMP-Fc
US-09-428-082B-403

  Query Match             62.7%;  Score 192;  DB 2;  Length 36;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-13;
  Matches   31;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         19 VCKPQGGGGGGGGTYSCHFGPLTWVCKPQGG 49
              |||||||||||||||||||||||||||||||
Db          1 VCKPQGGGGGGGGTYSCHFGPLTWVCKPQGG 31


RESULT 10
US-09-428-082B-16
; Sequence 16, Application US/09428082B
; Patent No. 6660843
; GENERAL INFORMATION:
;  APPLICANT: FEIGE, ULRICH
;  APPLICANT:  LIU, CHUAN-FA
;  APPLICANT:  CHEETHAM, JANET C.
;  APPLICANT:  BOONE, THOMAS CHARLES
;  TITLE OF INVENTION: MODIFIED PEPTIDES AS THERAPEUTIC AGENTS
;  FILE REFERENCE: A-527
;  CURRENT APPLICATION NUMBER: US/09/428,082B
;  CURRENT FILING DATE:  1999-10-22
;  PRIOR APPLICATION NUMBER: 60/105,371
;  PRIOR FILING DATE: 1998-10-23
;  NUMBER OF SEQ ID NOS: 1133
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16
;   LENGTH: 253
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Fc-EMP
US-09-428-082B-16

  Query Match             55.1%;  Score 168.5;  DB 2;  Length 253;
  Best Local Similarity   56.6%;  Pred. No. 2.8e-10;
  Matches   30;  Conservative    3;  Mismatches   11;  Indels    9;  Gaps    1;
```

```
Qy          6 GGTYSC---------HFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGG 49
              |  :||         |:   :    |  ||||||||||||||||||||||||
Db        201 GNVFSCSVMHEALHNHYTQKSLSLSPGKGGGGGGGTYSCHFGPLTWVCKPQGG 253
```

RESULT 11
US-09-428-082B-18
; Sequence 18, Application US/09428082B
; Patent No. 6660843
; GENERAL INFORMATION:
; APPLICANT: FEIGE, ULRICH
; APPLICANT: LIU, CHUAN-FA
; APPLICANT: CHEETHAM, JANET C.
; APPLICANT: BOONE, THOMAS CHARLES
; TITLE OF INVENTION: MODIFIED PEPTIDES AS THERAPEUTIC AGENTS
; FILE REFERENCE: A-527
; CURRENT APPLICATION NUMBER: US/09/428,082B
; CURRENT FILING DATE: 1999-10-22
; PRIOR APPLICATION NUMBER: 60/105,371
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 1133
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: EMP-Fc
US-09-428-082B-18

Query Match 52.8%; Score 161.5; DB 2; Length 253;
Best Local Similarity 87.1%; Pred. No. 1.5e-09;
Matches 27; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

```
Qy          6 GGTYSCHFGPLTWVCKPQGGGGGGGG-TYSC 35
              ||||||||||||||||||||||||  |::|
Db          2 GGTYSCHFGPLTWVCKPQGGGGGGGDKTHTC 32
```

RESULT 12
US-09-428-082B-1034
; Sequence 1034, Application US/09428082B
; Patent No. 6660843
; GENERAL INFORMATION:
; APPLICANT: FEIGE, ULRICH
; APPLICANT: LIU, CHUAN-FA
; APPLICANT: CHEETHAM, JANET C.
; APPLICANT: BOONE, THOMAS CHARLES
; TITLE OF INVENTION: MODIFIED PEPTIDES AS THERAPEUTIC AGENTS
; FILE REFERENCE: A-527
; CURRENT APPLICATION NUMBER: US/09/428,082B
; CURRENT FILING DATE: 1999-10-22
; PRIOR APPLICATION NUMBER: 60/105,371
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 1133
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1034
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: EPO-MIMETIC
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (25)..(25)
; OTHER INFORMATION: Fc domain attached at Position 25 of the C-terminus
US-09-428-082B-1034

Query Match 51.0%; Score 156; DB 2; Length 25;
Best Local Similarity 100.0%; Pred. No. 5.7e-10;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          6 GGTYSCHFGPLTWVCKPQGGGGGGG 30
              |||||||||||||||||||||||||
Db          1 GGTYSCHFGPLTWVCKPQGGGGGGG 25
```

RESULT 13
US-08-484-135-8
; Sequence 8, Application US/08484135
; Patent No. 5767078
; GENERAL INFORMATION:
; APPLICANT: Johnson, Dana L
; APPLICANT: Zivin, Robert A
; TITLE OF INVENTION: AGONIST PEPTIDE DIMERS
; NUMBER OF SEQUENCES: 93
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Frank S. DiGiglio
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: U.S.A..
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,135
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: DiGiglio, Frank S
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9594
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-484-135-8

Query Match 41.2%; Score 126; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 5.8e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          6 GGTYSCHFGPLTWVCKPQGG 25
              ||||||||||||||||||||
Db          1 GGTYSCHFGPLTWVCKPQGG 20
```

RESULT 14
US-08-484-135-20
; Sequence 20, Application US/08484135
; Patent No. 5767078
; GENERAL INFORMATION:
; APPLICANT: Johnson, Dana L
; APPLICANT: Zivin, Robert A
; TITLE OF INVENTION: AGONIST PEPTIDE DIMERS
; NUMBER OF SEQUENCES: 93
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Frank S. DiGiglio
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: U.S.A..
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/484,135
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DiGiglio, Frank S
;      REGISTRATION NUMBER:  31,346
;      REFERENCE/DOCKET NUMBER:  9594
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (516) 742-4343
;      TELEFAX:  (516) 742-4366
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-484-135-20

  Query Match             41.2%;  Score 126;  DB 1;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 5.8e-07;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 GGTYSCHFGPLTWVCKPQGG 25
              ||||||||||||||||||||
Db          1 GGTYSCHFGPLTWVCKPQGG 20


RESULT 15
US-08-484-135-42
; Sequence 42, Application US/08484135
; Patent No. 5767078
;  GENERAL INFORMATION:
;    APPLICANT:  Johnson, Dana L
;    APPLICANT:  Zivin, Robert A
;    TITLE OF INVENTION:  AGONIST PEPTIDE DIMERS
;    NUMBER OF SEQUENCES:  93
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Frank S. DiGiglio
;      STREET:  400 Garden City Plaza
;      CITY:  Garden City
;      STATE:  New York
;      COUNTRY:  U.S.A..
;      ZIP:  11530
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/484,135
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DiGiglio, Frank S
;      REGISTRATION NUMBER:  31,346
;      REFERENCE/DOCKET NUMBER:  9594
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (516) 742-4343
;      TELEFAX:  (516) 742-4366
;  INFORMATION FOR SEQ ID NO:  42:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-484-135-42

  Query Match             41.2%;  Score 126;  DB 1;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 5.8e-07;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 GGTYSCHFGPLTWVCKPQGG 25
              ||||||||||||||||||||
Db          1 GGTYSCHFGPLTWVCKPQGG 20


Search completed: March 31, 2006, 16:40:37
Job time : 49.0249 secs
```

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: March 31, 2006, 16:09:06 ; Search time 188.93 Seconds
(without alignments)
113.955 Million cell updates/sec

Title: US-10-609-217-340
Perfect score: 306
Sequence: 1 GGTYSCHFGPLTWVCKPQGG.....CHFGPLTWVCKPQGGGGGGG 49

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	306	100.0	49	5	ABB73393	Abb73393 EPO-mimet
2	306	100.0	50	3	AAB17284	Aab17284 EPO-mimet
3	306	100.0	57	3	AAB17314	Aab17314 EMP-EMP-F
4	306	100.0	57	5	ABB73408	Abb73408 EMP-EMP g
5	306	100.0	277	3	AAB16966	Aab16966 EMP-EMP-F
6	306	100.0	278	5	ABB73417	Abb73417 EMP-EMP-F
7	276	90.2	49	5	ABB73392	Abb73392 EPO-mimet
8	276	90.2	50	3	AAB17283	Aab17283 EPO-mimet
9	276	90.2	277	3	AAB16967	Aab16967 Fc-EMP-EM
10	276	90.2	277	5	ABB73418	Abb73418 Fc-EMP-EM
11	249	81.4	47	3	AAB17040	Aab17040 EPO-mimet
12	249	81.4	47	8	ADJ52198	Adj52198 CH1 delet
13	249	81.4	47	8	ADJ51160	Adj51160 CH1 delet
14	249	81.4	70	7	ADJ72562	Adj72562 EPO mimet
15	240	78.4	40	3	AAB17036	Aab17036 EPO-mimet
16	240	78.4	40	5	ABB72819	Abb72819 Erythropo
17	240	78.4	40	8	ADJ52195	Adj52195 CH1 delet
18	239.5	78.3	41	3	AAB17037	Aab17037 EPO-mimet
19	239.5	78.3	41	5	ABB72820	Abb72820 Erythropo
20	239.5	78.3	41	7	ADJ72559	Adj72559 EPO mimet
21	239.5	78.3	41	8	ADJ51157	Adj51157 CH1 delet
22	239.5	78.3	46	3	AAB17039	Aab17039 EPO-mimet
23	239.5	78.3	46	5	ABB72822	Abb72822 Erythropo
24	235	76.8	47	5	ABB72823	Abb72823 Erythropo
25	222	72.5	36	3	AAB17313	Aab17313 EMP-Fc fu
26	222	72.5	36	5	ABB73407	Abb73407 EPO mimet
27	206	67.3	145	7	ADJ73529	Adj73529 Erythropo
28	201	65.7	39	3	AAB17312	Aab17312 Fc-EMP fu
29	201	65.7	39	5	ABB73406	Abb73406 EPO mimet
30	168.5	55.1	253	3	AAB16964	Aab16964 Fc-EMP pr
31	168.5	55.1	253	5	ABB73415	Abb73415 Fc-EPO mi
32	161.5	52.8	253	3	AAB16965	Aab16965 EMP-Fc pr
33	161.5	52.8	253	5	ABB73416	Abb73416 EPO mimet
34	156	51.0	25	5	ABB73394	Abb73394 EPO-mimet
35	156	51.0	26	3	AAB17930	Aab17930 EPO-mimet
36	155.5	50.8	251	9	ADZ44485	Adz44485 Erythropo
37	152.5	49.8	266	8	ADJ52121	Adj52121 CH1 delet
38	152	49.7	51	8	ADJ52126	Adj52126 CH1 delet
39	152	49.7	51	8	ADJ52127	Adj52127 CH1 delet
40	152	49.7	269	8	ADJ52120	Adj52120 CH1 delet
41	150.5	49.2	37	8	ADJ52122	Adj52122 CH1 delet
42	150	49.0	129	7	ADJ73537	Adj73537 Erythropo
43	150	49.0	249	9	ADZ44490	Adz44490 Erythropo
44	150	49.0	249	9	ADZ44487	Adz44487 Erythropo
45	148	48.4	131	7	ADJ73539	Adj73539 Erythropo

ALIGNMENTS

RESULT 1
ABB73393
ID ABB73393 standard; peptide; 49 AA.
XX
AC ABB73393;
XX
DT 05-APR-2002 (first entry)
XX
DE EPO-mimetic peptide SEQ ID NO:340.
XX
KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200183525-A2.
XX
PD 08-NOV-2001.
XX
PF 02-MAY-2001; 2001WO-US014310.
XX
PR 03-MAY-2000; 2000US-00563286.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
XX
DR WPI; 2002-130313/17.
XX
PT Novel vehicle-peptide molecule or its multimers useful for treating
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT diabetic retinopathy, obesity, sleep disorders and infertility.
XX
PS Claim 16; Page 90; 176pp; English.
XX
CC The present invention describes a vehicle-peptide molecule (I) or its

CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive, cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological, antianaemic, anorectic, antiinfertility, haemostatic, dermatological and neuroprotective activities. (I) can be used as a therapeutic or prophylactic agent as well as for screening purposes. (I) is useful for diagnosing diseases characterised by dysfunction of their associated protein of interest, for identifying normal or abnormal proteins of interest, as a part of diagnostic kit to detect the presence of their proteins of interest in a biological sample. Additionally, (I) is useful for treating inflammatory and autoimmune diseases, tumour growth, cancer, rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders, infertility, and neurological degenerative diseases. (I), comprising EPO-mimetic compounds are useful for treating disorders characterised by low red blood cell levels such as anaemia. The TPO-mimetic comprising compounds are useful for treating conditions that involve an existing megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic tumour which result in thrombocytopaenia, systemic lupus erythematosus, and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777 represent amino acid and nucleic acid sequences used in the exemplification of the present invention
XX
SQ Sequence 49 AA;

```
  Query Match             100.0%;  Score 306;  DB 5;  Length 49;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-24;
  Matches   49;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGGGGGGG 49
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGGGGGGG 49
```

RESULT 2
AAB17284
ID AAB17284 standard; peptide; 50 AA.
XX
AC AAB17284;
XX
DT 31-OCT-2000 (first entry)
XX
DE EPO-mimetic peptide sequence SEQ ID NO:340.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US025044.
XX
PR 23-OCT-1998; 98US-0105371P.
PR 22-OCT-1999; 99US-00428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and pharmacologically
PT active peptides, useful for treating cancer and autoimmune diseases.
XX
PS Claim 16; Page 314; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an Fc domain, pharmacologically active peptides, and linkers. Where (I) is: (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-(L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2, P3, and P4 = are each independently sequences of pharmacologically active peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b, c, d, e, and f = are each independently 0 or 1, provided that at least 1 of a and b is 1. The composition can have cytostatic, antiasthmatic, thrombolytic and immunosuppressive activities. DNAs, vectors and host cells from the present invention can be used for producing pharmaceutical compositions. The compositions are useful for treating cancer, asthma, thrombosis, or autoimmune diseases. The use of an Fc domain (rather than a Fab domain) can provide a longer half-life or incorporate functions such as Fc receptor binding, protein A binding, complement fixation, and possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to AAB18003 represent nucleotide and amino acid sequences used in the exemplification of the present invention
XX
SQ Sequence 50 AA;

```
  Query Match             100.0%;  Score 306;  DB 3;  Length 50;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-24;
  Matches   49;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGGGGGGG 49
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGGGGGGG 49
```

RESULT 3
AAB17314
ID AAB17314 standard; peptide; 57 AA.
XX
AC AAB17314;
XX
DT 31-OCT-2000 (first entry)
XX
DE EMP-EMP-Fc fusion construction peptide SEQ ID NO:417.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US025044.
XX
PR 23-OCT-1998; 98US-0105371P.
PR 22-OCT-1999; 99US-00428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and pharmacologically
PT active peptides, useful for treating cancer and autoimmune diseases.
XX
PS Example 3; Page 342; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:

CC (X1)a-P1-(X2)b, where: P1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and
CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC AAB18003 represent nucleotide and amino acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 57 AA;

Query Match 100.0%; Score 306; DB 3; Length 57;
Best Local Similarity 100.0%; Pred. No. 2e-24;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGGGGGGG 49
Db 2 GGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGGGGGGG 50

RESULT 4
ABB73408
ID ABB73408 standard; peptide; 57 AA.
XX
AC ABB73408;
XX
DT 05-APR-2002 (first entry)
XX
DE EMP-EMP gene construction related peptide SEQ ID NO:417.
XX
KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200183525-A2.
XX
PD 08-NOV-2001.
XX
PF 02-MAY-2001; 2001WO-US014310.
XX
PR 03-MAY-2000; 2000US-00563286.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
XX
DR WPI; 2002-130313/17.
XX
PT Novel vehicle-peptide molecule or its multimers useful for treating
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT diabetic retinopathy, obesity, sleep disorders and infertility.
XX
PS Example 3; Page 116; 176pp; English.
XX
CC The present invention describes a vehicle-peptide molecule (I) or its
CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
CC neuroprotective activities. (I) can be used as a therapeutic or
CC prophylactic agent as well as for screening purposes. (I) is useful for
CC diagnosing diseases characterised by dysfunction of their associated
CC protein of interest, for identifying normal or abnormal proteins of
CC interest, as a part of diagnostic kit to detect the presence of their
CC proteins of interest in a biological sample. Additionally, (I) is useful
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
CC infertility, and neurological degenerative diseases. (I), comprising EPO-
CC mimetic compounds are useful for treating disorders characterised by low
CC red blood cell levels such as anaemia. The TPO-mimetic comprising
CC compounds are useful for treating conditions that involve an existing
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC represent amino acid and nucleic acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 57 AA;

Query Match 100.0%; Score 306; DB 5; Length 57;
Best Local Similarity 100.0%; Pred. No. 2e-24;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGGGGGGG 49
Db 2 GGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGGGGGGG 50

RESULT 5
AAB16966
ID AAB16966 standard; protein; 277 AA.
XX
AC AAB16966;
XX
DT 31-OCT-2000 (first entry)
XX
DE EMP-EMP-Fc protein sequence SEQ ID NO:20.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US025044.
XX
PR 23-OCT-1998; 98US-0105371P.
PR 22-OCT-1999; 99US-00428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
DR N-PSDB; AAA69450.
XX
PT Novel composition of matter comprising an Fc domain and pharmacologically

PT active peptides, useful for treating cancer and autoimmune diseases.
XX
PS Claim 17; Page 198-199; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and
CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC AAB18003 represent nucleotide and amino acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 277 AA;

Query Match 100.0%; Score 306; DB 3; Length 277;
Best Local Similarity 100.0%; Pred. No. 9.2e-24;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGGGGGGG 49
||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2 GGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGGGGGGG 50

RESULT 6
ABB73417
ID ABB73417 standard; protein; 278 AA.
XX
AC ABB73417;
XX
DT 05-APR-2002 (first entry)
XX
DE EMP-EMP-Fc amino acid SEQ ID NO:20.
XX
KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200183525-A2.
XX
PD 08-NOV-2001.
XX
PF 02-MAY-2001; 2001WO-US014310.
XX
PR 03-MAY-2000; 2000US-00563286.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
XX
DR WPI; 2002-130313/17.
DR N-PSDB; ABL35767.
XX
PT Novel vehicle-peptide molecule or its multimers useful for treating
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT diabetic retinopathy, obesity, sleep disorders and infertility.
XX
PS Claim 12; Fig 15; 176pp; English.
XX
CC The present invention describes a vehicle-peptide molecule (I) or its
CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
CC neuroprotective activities. (I) can be used as a therapeutic or
CC prophylactic agent as well as for screening purposes. (I) is useful for
CC diagnosing diseases characterised by dysfunction of their associated
CC protein of interest, for identifying normal or abnormal proteins of
CC interest, as a part of diagnostic kit to detect the presence of their
CC proteins of interest in a biological sample. Additionally, (I) is useful
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
CC infertility, and neurological degenerative diseases. (I), comprising EPO-
CC mimetic compounds are useful for treating disorders characterised by low
CC red blood cell levels such as anaemia. The TPO-mimetic comprising
CC compounds are useful for treating conditions that involve an existing
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC represent amino acid and nucleic acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 278 AA;

Query Match 100.0%; Score 306; DB 5; Length 278;
Best Local Similarity 100.0%; Pred. No. 9.2e-24;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGGGGGGG 49
||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2 GGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGGGGGGG 50

RESULT 7
ABB73392
ID ABB73392 standard; peptide; 49 AA.
XX
AC ABB73392;
XX
DT 05-APR-2002 (first entry)
XX
DE EPO-mimetic peptide SEQ ID NO:339.
XX
KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200183525-A2.
XX
PD 08-NOV-2001.
XX
PF 02-MAY-2001; 2001WO-US014310.

XX
PR 03-MAY-2000; 2000US-00563286.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
XX
DR WPI; 2002-130313/17.
XX
PT Novel vehicle-peptide molecule or its multimers useful for treating
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT diabetic retinopathy, obesity, sleep disorders and infertility.
XX
PS Claim 16; Page 90; 176pp; English.
XX
CC The present invention describes a vehicle-peptide molecule (I) or its
CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
CC neuroprotective activities. (I) can be used as a therapeutic or
CC prophylactic agent as well as for screening purposes. (I) is useful for
CC diagnosing diseases characterised by dysfunction of their associated
CC protein of interest, for identifying normal or abnormal proteins of
CC interest, as a part of diagnostic kit to detect the presence of their
CC proteins of interest in a biological sample. Additionally, (I) is useful
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
CC infertility, and neurological degenerative diseases. (I), comprising EPO-
CC mimetic compounds are useful for treating disorders characterised by low
CC red blood cell levels such as anaemia. The TPO-mimetic comprising
CC compounds are useful for treating conditions that involve an existing
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC represent amino acid and nucleic acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 49 AA;

Query Match 90.2%; Score 276; DB 5; Length 49;
Best Local Similarity 100.0%; Pred. No. 2.2e-21;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGG 44
||||||||||||||||||||||||||||||||||||||||||||
Db 6 GGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGG 49

RESULT 8
AAB17283
ID AAB17283 standard; peptide; 50 AA.
XX
AC AAB17283;
XX
DT 31-OCT-2000 (first entry)
XX
DE EPO-mimetic peptide sequence SEQ ID NO:339.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US025044.
XX
PR 23-OCT-1998; 98US-0105371P.
PR 22-OCT-1999; 99US-00428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and pharmacologically
PT active peptides, useful for treating cancer and autoimmune diseases.
XX
PS Claim 16; Page 314; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and
CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC AAB18003 represent nucleotide and amino acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 50 AA;

Query Match 90.2%; Score 276; DB 3; Length 50;
Best Local Similarity 100.0%; Pred. No. 2.2e-21;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGG 44
||||||||||||||||||||||||||||||||||||||||||||
Db 7 GGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGG 50

RESULT 9
AAB16967
ID AAB16967 standard; protein; 277 AA.
XX
AC AAB16967;
XX
DT 31-OCT-2000 (first entry)
XX
DE Fc-EMP-EMP protein sequence SEQ ID NO:22.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US025044.
XX

PR 23-OCT-1998; 98US-0105371P.
PR 22-OCT-1999; 99US-00428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
DR N-PSDB; AAA69451.
XX
PT Novel composition of matter comprising an Fc domain and pharmacologically
PT active peptides, useful for treating cancer and autoimmune diseases.
XX
PS Example 3; Page 201-202; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and
CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC AAB18003 represent nucleotide and amino acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 277 AA;

Query Match 90.2%; Score 276; DB 3; Length 277;
Best Local Similarity 100.0%; Pred. No. 1.2e-20;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGG 44
              ||||||||||||||||||||||||||||||||||||||||||||
Db        234 GGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGG 277
```

RESULT 10
ABB73418
ID ABB73418 standard; protein; 277 AA.
XX
AC ABB73418;
XX
DT 05-APR-2002 (first entry)
XX
DE Fc-EMP-EMP nucleic acid SEQ ID NO:22.
XX
KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG; EPO;
KW erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW antianaemic; anorectic; antiinfertility; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200183525-A2.
XX
PD 08-NOV-2001.
XX
PF 02-MAY-2001; 2001WO-US014310.
XX
PR 03-MAY-2000; 2000US-00563286.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
XX
DR WPI; 2002-130313/17.
DR N-PSDB; ABL35768.
XX
PT Novel vehicle-peptide molecule or its multimers useful for treating
PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
PT diabetic retinopathy, obesity, sleep disorders and infertility.
XX
PS Claim 12; Fig 16; 176pp; English.
XX
CC The present invention describes a vehicle-peptide molecule (I) or its
CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
CC antianaemic, anorectic, antiinfertility, haemostatic, dermatological and
CC neuroprotective activities. (I) can be used as a therapeutic or
CC prophylactic agent as well as for screening purposes. (I) is useful for
CC diagnosing diseases characterised by dysfunction of their associated
CC protein of interest, for identifying normal or abnormal proteins of
CC interest, as a part of diagnostic kit to detect the presence of their
CC proteins of interest in a biological sample. Additionally, (I) is useful
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
CC infertility, and neurological degenerative diseases. (I), comprising EPO-
CC mimetic compounds are useful for treating disorders characterised by low
CC red blood cell levels such as anaemia. The TPO-mimetic comprising
CC compounds are useful for treating conditions that involve an existing
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
CC deficiency, such as thrombocytopaenia, aplastic anaemia, metastatic
CC tumour which result in thrombocytopaenia, systemic lupus erythematosus,
CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC represent amino acid and nucleic acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 277 AA;

Query Match 90.2%; Score 276; DB 5; Length 277;
Best Local Similarity 100.0%; Pred. No. 1.2e-20;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGG 44
              ||||||||||||||||||||||||||||||||||||||||||||
Db        234 GGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGG 277
```

RESULT 11
AAB17040
ID AAB17040 standard; peptide; 47 AA.
XX
AC AAB17040;
XX
DT 31-OCT-2000 (first entry)
XX
DE EPO-mimetic peptide sequence SEQ ID NO:96.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX
OS Synthetic.
XX

PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US025044.
XX
PR 23-OCT-1998; 98US-0105371P.
PR 22-OCT-1999; 99US-00428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and pharmacologically
PT active peptides, useful for treating cancer and autoimmune diseases.
XX
PS Claim 13; Page 228; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and
CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC AAB18003 represent nucleotide and amino acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 47 AA;

Query Match 81.4%; Score 249; DB 3; Length 47;
Best Local Similarity 90.9%; Pred. No. 1.3e-18;
Matches 40; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGG 44
||||||||||||||||||| |||||||||||||||||||
Db 1 GGTYSCHFGPLTWVCKPQGGSSKXGGTYSCHFGPLTWVCKPQGG 44

RESULT 12
ADJ52198
ID ADJ52198 standard; peptide; 47 AA.
XX
AC ADJ52198;
XX
DT 06-MAY-2004 (first entry)
XX
DE CH1 deleted mimetibody-related peptide SeqID14.
XX
KW CH1 deleted mimetibody; immunosuppressive; cardiovascular; cardiant;
KW hypotensive; neuroprotective; nootropic; antibacterial; virucide;
KW fungicide; gene therapy; immune disorder; cardiovascular disease;
KW arrhythmia; hypertension; heart failure; neurodegenerative;
KW multiple sclerosis; dementia; Alzheimer's disease; anaemia;
KW cancerous condition; infectious disease; bacterial infection;
KW viral infection; fungal infection.
XX
OS Unidentified.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 24
FT /label= OTHER
FT /note= "OTHER= linker"
XX
PN WO2004002417-A2.
XX
PD 08-JAN-2004.
XX
PF 27-JUN-2003; 2003WO-US020347.
XX
PR 28-JUN-2002; 2002US-0392431P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Ghrayeb J, Scallon BJ, Nesspor TC;
PI Kutoloski KA;
XX
DR WPI; 2004-082870/08.
XX
PT New CH1-deleted mimetibody polypeptides and nucleic acids, useful for
PT modulating, treating, alleviating, preventing an immune, cardiovascular,
PT or neurodegenerative disease or disorder, anemia, cancer, or infectious
PT diseases.
XX
PS Example 1; SEQ ID NO 14; 129pp; English.
XX
CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an immunosuppressive,
CC cardiovascular, cardiant, hypotensive, neuroprotective, nootropic,
CC antibacterial, virucide or fungicide activity. In addition, the disclosed
CC sequences may prove useful for gene therapy. The CH1-deleted mimetibody
CC is useful for diagnosing or treating a disease condition in a cell,
CC tissue, organ or animal, specifically for modulating, treating,
CC alleviating, preventing the incidence or reducing the symptoms of an
CC immune, cardiovascular (for example arrhythmia, hypertension or heart
CC failure), or neurodegenerative (for example multiple sclerosis, dementia
CC or Alzheimer's disease) diseases or disorders, anaemia, cancerous
CC conditions, or infectious diseases (for example bacterial, viral or
CC fungal infection). The present sequence is that of a peptide which may be
CC used during the creation of a mimetibody of the invention.
XX
SQ Sequence 47 AA;

Query Match 81.4%; Score 249; DB 8; Length 47;
Best Local Similarity 90.9%; Pred. No. 1.3e-18;
Matches 40; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGG 44
||||||||||||||||||| |||||||||||||||||||
Db 1 GGTYSCHFGPLTWVCKPQGGSSKXGGTYSCHFGPLTWVCKPQGG 44

RESULT 13
ADJ51160
ID ADJ51160 standard; peptide; 47 AA.
XX
AC ADJ51160;
XX
DT 06-MAY-2004 (first entry)
XX
DE CH1 deleted mimetibody-related peptide SeqID14.
XX
KW CH1 deleted mimetibody; osteopathic; cardiovascular-Gen;
KW dermatological-Gen; auditory; endocrine-Gen; gastrointestinal-Gen;
KW gynaecological-Gen; hepatotropic; haemostatic; immunomodulator;
KW antiallergic; muscular-Gen; cytostatic; antiinflammatory; neuroleptic;
KW ophthalmological; nephrotropic; respiratory-Gen; tumour necrosis factor;
KW TNF; cytokine; bone disorder; joint disorder; cardiovascular disorder;
KW dental disorder; oral disorder; dermatological disorder; ear disorder;
KW nose disorder; throat disorder; endocrine disorder; metabolic disorder;
KW gastrointestinal disorder; gynaecological disorder; hepatic disorder;
KW obstetric disorder; haematologic disorder; immunological disorder;
KW allergic disorder; infectious disorder; musculoskeletal disorder;

KW oncological disorder; neurological disorder; nutritional disorder;
KW ophthalmologic disorder; pediatric disorder; psychiatric disorder;
KW renal disorder; pulmonary disorder.
XX
OS Unidentified.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 24
FT /label= OTHER
FT /note= "OTHER= linker"
XX
PN WO2004002424-A2.
XX
PD 08-JAN-2004.
XX
PF 30-JUN-2003; 2003WO-US020495.
XX
PR 28-JUN-2002; 2002US-0392431P.
PR 19-SEP-2002; 2002US-0412144P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Ghrayeb J, Scallon BJ, Nesspor TC;
PI Kutoloski KA;
XX
DR WPI; 2004-082872/08.
XX
PT New CH1 deleted mimetibody polypeptide and nucleic acid, useful for
PT diagnosing, preventing or treating cardiovascular, dermatologic,
PT endocrine, gastrointestinal, gynecologic, infectious, neurologic and
PT nutritional disorders.
XX
PS Claim 8; SEQ ID NO 14; 123pp; English.
XX
CC This invention relates to CH1 deleted mimetibodies (and the DNA sequences
CC which encode them), compositions, methods and uses. The invention may be
CC useful for the development of compounds with an osteopathic,
CC cardiovascular-Gen, dermatological-Gen, auditory, endocrine-Gen,
CC gastrointestinal-Gen, gynaecological-Gen, hepatotropic, haemostatic,
CC immunomodulator, antiallergic, muscular-Gen, cytostatic,
CC antiinflammatory, neuroleptic, ophthalmological, nephrotropic or
CC respiratory-Gen activity acting as a tumour necrosis factor (TNF)-
CC modulator or cytokine-agonist. The methods and compositions of the
CC present invention are useful for the diagnosis, prevention and/or
CC treatment of diseases or conditions associated with aberrant expression
CC or activity of the CH1 deleted mimetibody, such as a bone or joint,
CC cardiovascular, dental or oral, dermatological, ear, nose or throat,
CC endocrine, metabolic, gastrointestinal, gynaecological, hepatic,
CC obstetric, haematologic, immunological, allergic, infectious,
CC musculoskeletal, oncological, neurological, nutritional, ophthalmologic,
CC pediatric, psychiatric, renal or pulmonary disorders. The present
CC sequence is that of a peptide which may be used during the creation of a
CC mimetibody of the invention.
XX
SQ Sequence 47 AA;

Query Match 81.4%; Score 249; DB 8; Length 47;
Best Local Similarity 90.9%; Pred. No. 1.3e-18;
Matches 40; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGG 44
||||||||||||||||||| |||||||||||||||||||
Db 1 GGTYSCHFGPLTWVCKPQGGSSKXGGTYSCHFGPLTWVCKPQGG 44

RESULT 14
ADJ72562
ID ADJ72562 standard; peptide; 70 AA.
XX
AC ADJ72562;
XX
DT 06-MAY-2004 (first entry)
XX
DE EPO mimetic peptide sequence SeqID 14.
XX
KW mimetic; CDR mimetibody; gene therapy; transgenic; immune;
KW cardiovascular; infectious; malignant; neurologic disease; anaemia;
KW immunomodulator; cardiant; antimicrobial; cytostatic; neuroprotective;
KW erythropoietin; EPO.
XX
OS Synthetic.
XX
PN WO2003084477-A2.
XX
PD 16-OCT-2003.
XX
PF 24-MAR-2003; 2003WO-US009139.
XX
PR 29-MAR-2002; 2002US-0368791P.
XX
PA (CENZ) CENTOCOR INC.
XX
PI Heavner GA, Knight DM, Scallon BJ, Ghrayeb J;
XX
DR WPI; 2003-804237/75.
XX
PT New CDR mimetibody comprising a portion of a heavy or light chain
PT variable region comprising human framework or ligand binding region,
PT useful for preparing a composition for treating e.g., immune,
PT cardiovascular or neurologic disease.
XX
PS Disclosure; SEQ ID NO 14; 97pp; English.
XX
CC This invention relates to novel mammalian CDR mimetibodies, specific
CC portions or variants thereof. Specifically, it refers to an antibody
CC fragment where a protein has been inserted into, or replaces a portion
CC of, one or more CDR regions, such that each CDR mimetibody comprises at
CC least one portion of a heavy chain or light chain variable region, which
CC itself comprises at least one human framework region and at least one
CC ligand binding region (LBR). The present invention describes human
CC mimetibodies, including modified immunoglobulins and cleavage products
CC that can be useful in gene therapy and the generation of transgenic
CC plants and animals. Furthermore, the CDR mimetibody is useful for
CC preparing compositions for modulating, treating or reducing the symptoms
CC of immune, cardiovascular, infectious, malignant and/ or neurologic
CC diseases, as well as anaemia. Accordingly, they exhibit immunomodulator,
CC cardiant, antimicrobial, cytostatic and neuroprotective activities. This
CC peptide sequence is a erythropoietin (EPO) mimetic peptide sequence used
CC to make a mimetibody of the invention.
XX
SQ Sequence 70 AA;

Query Match 81.4%; Score 249; DB 7; Length 70;
Best Local Similarity 90.9%; Pred. No. 1.9e-18;
Matches 40; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGG 44
||||||||||||||||||| |||||||||||||||||||
Db 24 GGTYSCHFGPLTWVCKPQGGSSKXGGTYSCHFGPLTWVCKPQGG 67

RESULT 15
AAB17036
ID AAB17036 standard; peptide; 40 AA.
XX
AC AAB17036;
XX
DT 31-OCT-2000 (first entry)
XX
DE EPO-mimetic peptide sequence SEQ ID NO:92.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;

KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US025044.
XX
PR 23-OCT-1998; 98US-0105371P.
PR 22-OCT-1999; 99US-00428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and pharmacologically
PT active peptides, useful for treating cancer and autoimmune diseases.
XX
PS Claim 13; Page 226; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and
CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC AAB18003 represent nucleotide and amino acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 40 AA;

Query Match 78.4%; Score 240; DB 3; Length 40;
Best Local Similarity 90.9%; Pred. No. 9.3e-18;
Matches 40; Conservative 0; Mismatches 0; Indels 4; Gaps 1;

```
Qy          1 GGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGG 44
              |||||||||||||||||    ||||||||||||||||||||||
Db          1 GGTYSCHFGPLTWVCKPQ----GGGGTYSCHFGPLTWVCKPQGG 40
```

Search completed: March 31, 2006, 16:22:26
Job time : 188.93 secs

OM protein - protein search, using sw model

Run on: March 31, 2006, 16:22:51 ; Search time 30.4726 Seconds
(without alignments)
154.717 Million cell updates/sec

Title: US-10-609-217-340
Perfect score: 306
Sequence: 1 GGTYSCHFGPLTWVCKPQGG.....CHFGPLTWVCKPQGGGGGGG 49

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	91	29.7	299	2	T00837	glycine-rich prote
2	89	29.1	1585	2	T31611	hypothetical prote
3	83.5	27.3	1276	2	E96776	hypothetical prote
4	83	27.1	201	2	F84596	glycine-rich prote
5	82.5	27.0	200	2	S10334	glycine-rich prote
6	82.5	27.0	214	1	KNNT2S	glycine-rich prote
7	82.5	27.0	343	2	T29547	hypothetical prote
8	82	26.8	280	2	A42424	chitinase (EC 3.2.
9	81	26.5	481	2	A35628	loricrin - mouse
10	81	26.5	892	2	T27005	hypothetical prote
11	80.5	26.3	146	2	T06796	glycine-rich RNA-b
12	80.5	26.3	228	2	D86416	probable beta-1,3
13	80	26.1	207	2	B44994	eggshell protein 2
14	80	26.1	212	2	A44994	eggshell protein 1
15	79.5	26.0	280	2	G84839	late embryogenesis
16	79.5	26.0	369	1	TVFVAF	transforming prote
17	79.5	26.0	653	1	S44749	C06G4.2 protein -
18	79	25.8	271	2	S34666	glycine-rich prote
19	79	25.8	420	2	A49642	transcription fact
20	79	25.8	919	2	A39248	androgen receptor
21	78.5	25.7	171	2	H84709	probable glycine-r
22	78	25.5	239	2	S49193	GCR 101 protein -
23	78	25.5	694	2	F70868	hypothetical glyci
24	77.5	25.3	165	1	KNRZG1	glycine-rich cell
25	77.5	25.3	385	2	T20410	hypothetical prote
26	77	25.2	404	2	S54729	RNA-binding protei
27	76.5	25.0	183	2	PN0109	keratin-like prote
28	76.5	25.0	224	2	T51742	RNA helicase RH11
29	76.5	25.0	299	2	T05494	glycine-rich prote
30	76.5	25.0	305	2	T20906	hypothetical prote
31	76.5	25.0	340	2	T20807	hypothetical prote
32	76.5	25.0	431	1	WJHU2G	homeotic protein H
33	76.5	25.0	482	2	T48337	hypothetical prote
34	76.5	25.0	603	2	T45671	ATP-dependent RNA
35	76	24.8	188	2	S49192	GCR 1 protein - fr
36	76	24.8	220	2	A44805	eggshell protein p
37	76	24.8	221	2	T04592	glycine-rich cell
38	76	24.8	285	2	T31503	hypothetical prote
39	76	24.8	336	1	S18750	chitinase (EC 3.2.
40	76	24.8	411	2	A49127	homeotic protein A
41	76	24.8	910	2	A34721	androgen receptor
42	76	24.8	911	2	B34721	androgen receptor
43	75.5	24.7	172	2	D41132	collagen-related p
44	75.5	24.7	312	2	T25048	hypothetical prote
45	75	24.5	255	2	B84777	hypothetical prote

ALIGNMENTS

RESULT 1
T00837
glycine-rich protein T13L16.11 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 31-Dec-2004
C;Accession: T00837; D84557
R;de la Bastide, M.; Hameed, A.; Gnoj, L.; Jensen, K.; Shohdy, N.; Gottesman, T.; Haberma
McCombie, W.R.
submitted to the EMBL Data Library, January 1999
A;Description: A. thaliana BAC T13L16 from chromosome IV, top arm.
A;Reference number: Z14205
A;Accession: T00837
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-299 <DEL>
A;Cross-references: UNIPROT:Q94C69; UNIPARC:UPI0000177E58; EMBL:AC003952; NID:g2708736;
A;Experimental source: cultivar Columbia
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: D84557
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-299 <STO>
A;Cross-references: UNIPARC:UPI0000177E58; GB:AE002093; NID:g2708747; PIDN:AAD03571.1;
C;Genetics:
A;Gene: At2g17870; T13L16.11
A;Map position: 2
C;Superfamily: cold shock domain homology
F;11-71/Domain: cold shock domain homology <CSD>

Query Match 29.7%; Score 91; DB 2; Length 299;
Best Local Similarity 44.8%; Pred. No. 0.035;
Matches 26; Conservative 3; Mismatches 15; Indels 14; Gaps 6;

```
Qy          2 GTYSC----HFGPLTWVCKPQGGG--GGGGGT-YSC-HFGPLTWVCK---PQGGGGGG 48
              | | |    ||      |: |||  |||| | |:|   | :  ||    | ||||||
Db        195 GCYMCGGVGHFAR---DCRQNGGGNVGGGGSTCYTCGGVGHIAKVCTSKIPSGGGGGG 249
```

RESULT 2
T31611
hypothetical protein Y50E8A.g - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C;Accession: T31611
R;Steward, C.
submitted to the EMBL Data Library, September 1999
A;Reference number: Z21047

```
A;Accession: T31611
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1585 <WIL>
A;Cross-references: UNIPARC:UPI000017BC9F; EMBL:AL117200; NID:e1549770; PIDN:CAB55050.1;
A;Experimental source: clone Y50E8A
C;Genetics:
A;Gene: CESP:Y50E8A.g
A;Introns: 25/3; 60/1; 133/2; 217/3; 270/3; 337/2; 400/1; 746/2

  Query Match             29.1%;  Score 89;  DB 2;  Length 1585;
  Best Local Similarity   40.8%;  Pred. No. 0.25;
  Matches   20;  Conservative    2;  Mismatches    5;  Indels   22;  Gaps    2;

Qy          1 GGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGGGGGGG 49
              || |:           |||||  || |:          || |||||||
Db        456 GGGYA-----------SGGGGGAGGGYA----------KPSGGGGGGG 482


RESULT 3
E96776
hypothetical protein F25A4.30 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: E96776
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: E96776
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1276 <STO>
A;Cross-references: UNIPROT:Q9SSF7; UNIPARC:UPI00000A6592; GB:AE005173; NID:g5882720; PI
C;Genetics:
A;Gene: F25A4.30
A;Map position: 1

  Query Match             27.3%;  Score 83.5;  DB 2;  Length 1276;
  Best Local Similarity   46.5%;  Pred. No. 0.7;
  Matches   20;  Conservative    2;  Mismatches    6;  Indels   15;  Gaps    3;

Qy          8 FGPLTW-VCKPQGGGGGGGGTYSCHFGPLTWVCKPQGGGGGGG 49
              | | |   : :||||||||      ||        |||||||
Db        599 FVPSKWHTLEGEGGGGGGGG------GP--------GGGGGGG 627


RESULT 4
F84596
glycine-rich protein (AtGRP2) [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: F84596
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: F84596
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-201 <STO>
A;Cross-references: UNIPROT:Q38896; UNIPARC:UPI0000000E38; GB:AE002093; NID:g4803937; PI
C;Genetics:
A;Gene: At2g21060
A;Map position: 2
C;Superfamily: Arabidopsis glycine-rich protein 2; cold shock domain homology

  Query Match             27.1%;  Score 83;  DB 2;  Length 201;
  Best Local Similarity   41.1%;  Pred. No. 0.15;
  Matches   23;  Conservative    1;  Mismatches   16;  Indels   16;  Gaps    3;

Qy          1 GGTYSC----HFGPLTWVCKPQGG---GGGGGGTYSCHFGPLTWVCKPQGGGGGGG 49
              ||  ||      | :   |   ||   |||||| |    |         |||||||
Db        133 GGDNSCFKCGEPGHMARECSQGGGGYSGGGGGGRYGSGGG---------GGGGGGG 179


RESULT 5
S10334
glycine-rich protein precursor - barley
C;Species: Hordeum vulgare (barley)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C;Accession: S10334
R;Rohde, W.; Rosch, K.; Kroeger, K.; Salamini, F.
Plant Mol. Biol. 14, 1057-1059, 1990
A;Title: Nucleotide sequence of a Hordeum vulgare gene encoding a glycine-rich protein wi
A;Reference number: S10334; MUID:91346692; PMID:1715208
A;Accession: S10334
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-200 <ROH>
A;Cross-references: UNIPROT:P17816; UNIPARC:UPI000012BB4B; EMBL:X52580; NID:g18995; PIDN:
C;Genetics:
A;Introns: 29/1
C;Superfamily: Arabidopsis glycine-rich protein 3

  Query Match             27.0%;  Score 82.5;  DB 2;  Length 200;
  Best Local Similarity   44.0%;  Pred. No. 0.17;
  Matches   22;  Conservative    2;  Mismatches    9;  Indels   17;  Gaps    3;

Qy          1 GGTYSCHFGPLTWVCKPQGGGG-GGGGTYSCHFGPLTWVCKPQGGGGGGG 49
              || |  | |          :|||| |||| |  | |          :|||| ||
Db         86 GGGYPGHGG--------EGGGGYGGGGGYPGHGG--------EGGGGYGG 119


RESULT 6
KNNT2S
glycine-rich protein 2 - wood tobacco
C;Species: Nicotiana sylvestris (wood tobacco)
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
C;Accession: S17731
R;Obokata, J.; Ohme, M.; Hayashida, N.
Plant Mol. Biol. 17, 953-955, 1991
A;Title: Nucleotide sequence of a cDNA clone encoding a putative glycine-rich protein of
A;Reference number: S17731; MUID:92003709; PMID:1912512
A;Accession: S17731
A;Molecule type: mRNA
A;Residues: 1-214 <OBO>
A;Cross-references: UNIPROT:P27484; UNIPARC:UPI000012BB03; EMBL:X60007; NID:g19742; PIDN
C;Superfamily: Arabidopsis glycine-rich protein 2; cold shock domain homology
C;Keywords: zinc finger
F;11-71/Domain: cold shock domain homology <CSD>
F;82-158/Region: glycine-rich
F;159-172/Region: zinc finger CCHC motif
F;176-195/Region: glycine-rich
F;196-209/Region: zinc finger CCHC motif

  Query Match             27.0%;  Score 82.5;  DB 1;  Length 214;
  Best Local Similarity   50.0%;  Pred. No. 0.18;
  Matches   16;  Conservative    1;  Mismatches   14;  Indels    1;  Gaps    1;

Qy         19 GGGGGGGGTYSC-HFGPLTWVCKPQGGGGGGG 49
              || ||| | : |   |      |   |||||||
Db        151 GGSGGGSGCFKCGESGHFARDCSQSGGGGGGG 182


RESULT 7
T29547
```

```
hypothetical protein F48C1.8 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T29547
R;Gattung, S.; Le, T.T.
submitted to the EMBL Data Library, April 1997
A;Description: The sequence of C. elegans cosmid F48C1.
A;Reference number: Z20638
A;Accession: T29547
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-343 <GAT>
A;Cross-references: UNIPROT:O01575; UNIPARC:UPI0000082610; EMBL:U97015; PIDN:AAB52349.1;
A;Experimental source: strain Bristol N2; clone F48C1
C;Genetics:
A;Gene: CESP:F48C1.8
A;Map position: 1
A;Introns: 16/1; 95/2; 132/1; 166/1

  Query Match             27.0%;  Score 82.5;  DB 2;  Length 343;
  Best Local Similarity   51.4%;  Pred. No. 0.27;
  Matches   19;  Conservative    1;  Mismatches    6;  Indels   11;  Gaps    2;

Qy         15 CKPQGGG--GGGGGTYSCHFGPLTWVCKPQGGGGGGG 49
              |:  |||  |||||  | | |          |||||||
Db        169 CEAHGGGHHGGGGGGGSHHHG---------GGGGGGG 196


RESULT 8
A42424
chitinase (EC 3.2.1.14) A - maize
C;Species: Zea mays (maize)
C;Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 17-Mar-1999
C;Accession: A42424; A42260
R;Huynh, Q.K.; Hironaka, C.M.; Levine, E.B.; Smith, C.E.; Borgmeyer, J.R.; Shah, D.M.
J. Biol. Chem. 267, 6635-6640, 1992
A;Title: Antifungal proteins from plants. Purification, molecular cloning, and antifunga
A;Reference number: A42424; MUID:92202208; PMID:1551872
A;Accession: A42424
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-280 <HUY>
A;Cross-references: UNIPARC:UPI0000175A96
A;Experimental source: seed
A;Note: sequence inconsistent with nucleotide translation
A;Note: sequence extracted from NCBI backbone (NCBIN:89874, NCBIP:89876)
R;Verburg, J.G.; Smith, C.E.; Lisek, C.A.; Huynh, Q.K.
J. Biol. Chem. 267, 3886-3893, 1992
A;Title: Identification of an essential tyrosine residue in the catalytic site of a chit
opyl)- carbodiimide.
A;Reference number: A42260; MUID:92156129; PMID:1740436
A;Accession: A42260
A;Molecule type: protein
A;Residues: 180-195 <VER>
A;Cross-references: UNIPARC:UPI0000175A97
A;Note: the residue designated 'X' was determined to be derivatized tyrosine; therefore,
C;Superfamily: lectin-related plant chitinase; hevein chitin-binding domain homology; pl
C;Keywords: glycosidase; hydrolase; polysaccharide degradation
F;26-61/Domain: hevein chitin-binding domain homology <HCB>
F;82-280/Domain: plant chitinase homology <PCH>
F;188/Active site: Tyr #status predicted

  Query Match             26.8%;  Score 82;  DB 2;  Length 280;
  Best Local Similarity   40.9%;  Pred. No. 0.26;
  Matches   18;  Conservative    1;  Mismatches    5;  Indels   20;  Gaps    2;

Qy          6 CHFGPLTWVCKPQGGGGGGGTYSCHFGPLTWVCKPQGGGGGGG 49
              |  ||    |: ||||||||               |||||  |
Db         53 CQSGP----CRSGGGGGGGG---------------GGGGGSG 76


RESULT 9
A35628
loricrin - mouse
C;Species: Mus musculus (house mouse)
C;Date: 21-Sep-1990 #sequence_revision 21-Sep-1990 #text_change 09-Jul-2004
C;Accession: A35628
R;Mehrel, T.; Hohl, D.; Rothnagel, J.A.; Longley, M.A.; Bundman, D.; Cheng, C.; Lichti, U
Cell 61, 1103-1112, 1990
A;Title: Identification of a major keratinocyte cell envelope protein, loricrin.
A;Reference number: A35628; MUID:90275605; PMID:2190691
A;Accession: A35628
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-481 <MEH>
A;Cross-references: UNIPROT:P18165; UNIPARC:UPI00000272CA; GB:M34398; NID:g198870; PIDN:A
C;Superfamily: loricrin
C;Keywords: cornified cell envelope; epidermis

  Query Match             26.5%;  Score 81;  DB 2;  Length 481;
  Best Local Similarity   42.9%;  Pred. No. 0.52;
  Matches   21;  Conservative    1;  Mismatches   19;  Indels    8;  Gaps    1;

Qy          1 GGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGGGGGGG 49
              ||  ||  |          ||||||| ||   |  :  |      |||||
Db        370 GGGGSCGGG--------SSGGGGGGGCYSSGGGGSSGGCGGGYSGGGGG 410


RESULT 10
T27005
hypothetical protein Y48B6A.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T27005
R;Wall, M.
submitted to the EMBL Data Library, September 1999
A;Reference number: Z20297
A;Accession: T27005
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-892 <WIL>
A;Cross-references: UNIPROT:Q9U299; UNIPARC:UPI0000075372; EMBL:AL110490; NID:e1542263; P
A;Experimental source: clone Y48B6A
C;Genetics:
A;Gene: CESP:Y48B6A.3
A;Introns: 96/3; 193/3; 419/1; 550/2; 691/3

  Query Match             26.5%;  Score 81;  DB 2;  Length 892;
  Best Local Similarity   35.5%;  Pred. No. 0.9;
  Matches   22;  Conservative    4;  Mismatches   18;  Indels   18;  Gaps    2;

Qy          1 GGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHF------------GPLTWVCKPQGGGGG 47
              || |   :|          ||||||||:|   :             ||  :   |  ||||
Db        828 GGGYGGGYGG-----GGGGGGGGGGGSYHQPYNQDQRRGGRGGGGGPPGYQRPPYRGGGG 882

Qy         48 GG 49
              ||
Db        883 GG 884


RESULT 11
T06796
glycine-rich RNA-binding protein - garden pea
C;Species: Pisum sativum (garden pea)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C;Accession: T06796
R;LaRosa, T.J.; Watson, J.C.
submitted to the EMBL Data Library, December 1996
A;Reference number: Z15821
A;Accession: T06796
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-146 <LAR>
A;Cross-references: UNIPROT:P93486; UNIPARC:UPI00000AC7ED; EMBL:U81287; NID:g1778373; PI
A;Experimental source: cv. Alaska
C;Superfamily: glycine-rich RNA-binding protein; ribonucleoprotein repeat homology
```

F;37-104/Domain: ribonucleoprotein repeat homology <RRM>

Query Match 26.3%; Score 80.5; DB 2; Length 146;
Best Local Similarity 51.5%; Pred. No. 0.2;
Matches 17; Conservative 2; Mismatches 7; Indels 7; Gaps 1;

```
Qy         17 PQGGGGGGGGTYSCHFGPLTWVCKPQGGGGGGG 49
              |:||||| || |   |         :|||||||
Db        119 PRGGGGGYGGGYGDGF-------SNRGGGGGG 144
```

RESULT 12
D86416
probable beta-1,3 glucanase, 26636-27432 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: D86416
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: D86416
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-228 <STO>
A;Cross-references: UNIPROT:Q9C7R3; UNIPARC:UPI00000482EC; GB:AE005172; NID:g10092219; P
C;Genetics:
A;Map position: 1

Query Match 26.3%; Score 80.5; DB 2; Length 228;
Best Local Similarity 45.1%; Pred. No. 0.3;
Matches 23; Conservative 3; Mismatches 14; Indels 11; Gaps 3;

```
Qy          1 GGTYSCHFGPLTWVCKPQGG--GGGGGGTYSCHFGPLTWVCKPQGGGGGGG 49
              | | :  : ||     | |||  ||||||     :|  |    | |||||||
Db         79 GTTPTGGYPPLYGTTPPGGGDVGGGGGG-----YGGGT----PGGGGGGGG 120
```

RESULT 13
B44994
eggshell protein 2A precursor - fluke (Schistosoma japonicum)
C;Species: Schistosoma japonicum
C;Date: 14-May-1993 #sequence_revision 14-May-1993 #text_change 09-Jul-2004
C;Accession: B44994
R;Henkle, K.J.; Cook, G.A.; Foster, L.A.; Engman, D.M.; Bobek, L.A.; Cain, G.D.; Donelso
Mol. Biochem. Parasitol. 42, 69-82, 1990
A;Title: The gene family encoding eggshell proteins of Schistosoma japonicum.
A;Reference number: A44994; MUID:91042837; PMID:2172818
A;Accession: B44994
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-207 <HEN>
A;Cross-references: UNIPROT:P19469; UNIPARC:UPI0000129DF7; GB:M32280; NID:g160974; PID:g
C;Superfamily: fluke eggshell protein
C;Keywords: egg shell
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-207/Product: eggshell protein #status predicted <MAT>

Query Match 26.1%; Score 80; DB 2; Length 207;
Best Local Similarity 47.4%; Pred. No. 0.31;
Matches 18; Conservative 1; Mismatches 5; Indels 14; Gaps 2;

```
Qy         19 GGGGGGGGTY-------SCHFGPLTWVCKPQGGGGGGG 49
              |||||||| |       |: |         |||||||
Db         32 GGGGGGGGGYGGWCGGSDCYGG-------GNGGGGGGG 62
```

RESULT 14
A44994
eggshell protein 1 precursor - fluke (Schistosoma japonicum)
C;Species: Schistosoma japonicum
C;Date: 14-May-1993 #sequence_revision 14-May-1993 #text_change 09-Jul-2004
C;Accession: A44994; A61146
R;Henkle, K.J.; Cook, G.A.; Foster, L.A.; Engman, D.M.; Bobek, L.A.; Cain, G.D.; Donelson
Mol. Biochem. Parasitol. 42, 69-82, 1990
A;Title: The gene family encoding eggshell proteins of Schistosoma japonicum.
A;Reference number: A44994; MUID:91042837; PMID:2172818
A;Accession: A44994
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-212 <HEN>
A;Cross-references: UNIPROT:P19470; UNIPARC:UPI0000129DF5; GB:M32281; NID:g160972; PID:gJ
R;Bobek, L.A.; Rekosh, D.M.; LoVerde, P.T.
Exp. Parasitol. 72, 381-390, 1991
A;Title: Schistosoma japonicum: analysis of eggshell protein genes, their expression, and
A;Reference number: A61146; MUID:91224244; PMID:1709112
A;Accession: A61146
A;Molecule type: DNA
A;Residues: 1-21,'S',23-84,90-212 <BOB>
A;Cross-references: UNIPARC:UPI000007A98D; GB:M59318; NID:g160967; PID:g160968
C;Superfamily: fluke eggshell protein
C;Keywords: egg shell
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-212/Product: eggshell protein #status predicted <MAT>

Query Match 26.1%; Score 80; DB 2; Length 212;
Best Local Similarity 47.4%; Pred. No. 0.31;
Matches 18; Conservative 1; Mismatches 5; Indels 14; Gaps 2;

```
Qy         19 GGGGGGGGTY-------SCHFGPLTWVCKPQGGGGGGG 49
              |||||||| |       |: |         |||||||
Db         32 GGGGGGGGGYGGWCGGSDCYGG-------GNGGGGGGG 62
```

RESULT 15
G84839
late embryogenesis abundant M17 protein [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: G84839
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: G84839
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-280 <STO>
A;Cross-references: UNIPROT:Q9S7S3; UNIPARC:UPI000009DE6F; GB:AE002093; NID:g3894196; PI
C;Genetics:
A;Gene: At2g41260
A;Map position: 2

Query Match 26.0%; Score 79.5; DB 2; Length 280;
Best Local Similarity 30.4%; Pred. No. 0.45;
Matches 21; Conservative 4; Mismatches 9; Indels 35; Gaps 3;

```
Qy         16 KPQ-----GGGGGGGGTYSCHF--------GPLTWVCKPQ---------- 42
              :||     ||||||||   | :        |   : |: |
Db        114 EPQQGGRGGGGGGGGGRGGCRWGCCGGWWRGRCRYCCRSQAEASEVVETVEPNDVEPQQG 173

Qy         43 --GGGGGGG 49
                |||||||
Db        174 GRGGGGGGG 182
```

Search completed: March 31, 2006, 16:37:19

Job time : 32.4726 secs

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: March 31, 2006, 16:09:36 ; Search time 183.567 Seconds
(without alignments)
188.328 Million cell updates/sec

Title: US-10-609-217-340
Perfect score: 306
Sequence: 1 GGTYSCHFGPLTWVCKPQGG.....CHFGPLTWVCKPQGGGGGGG 49

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	113	36.9	775	2	Q4W7T9_9CRUS	Q4w7t9 daphnia mag
2	113	36.9	779	2	Q4R1V0_9CRUS	Q4r1v0 daphnia mag
3	101.5	33.2	241	2	Q6YUR8_ORYSA	Q6yur8 oryza sativ
4	97	31.7	229	2	Q8LPA7_WHEAT	Q8lpa7 triticum ae
5	95.5	31.2	231	2	Q75QN8_WHEAT	Q75qn8 triticum ae
6	94	30.7	210	2	Q528N5_MAGGR	Q528n5 magnaporthe
7	91.5	29.9	688	2	Q9GNP2_CIOSA	Q9gnp2 ciona savig
8	91.5	29.9	770	2	Q9GNP1_CIOSA	Q9gnp1 ciona savig
9	91.5	29.9	797	2	Q9GV13_HYDMA	Q9gv13 hydra magni
10	91	29.7	299	2	Q7XJR8_ARATH	Q7xjr8 arabidopsis
11	91	29.7	301	2	Q94C69_ARATH	Q94c69 arabidopsis
12	89.5	29.2	410	2	Q6YUX6_ORYSA	Q6yux6 oryza sativ
13	89.5	29.2	472	2	Q7YXK1_ASCSU	Q7yxk1 ascaris suu
14	89	29.1	117	2	Q9VD49_DROME	Q9vd49 drosophila
15	88.5	28.9	362	1	NKX23_MOUSE	P97334 mus musculu
16	88.5	28.9	473	2	Q84WQ1_ARATH	Q84wq1 arabidopsis
17	88	28.8	101	2	Q86DK6_9ECHO	Q86dk6 echinometra
18	88	28.8	305	1	ROA0_HUMAN	Q13151 homo sapien
19	88	28.8	305	2	Q6IB18_HUMAN	Q6ib18 homo sapien
20	88	28.8	324	2	O35295_MOUSE	O35295 mus musculu
21	87	28.4	106	2	Q24921_9ECHO	Q24921 echinometra
22	87	28.4	205	2	Q75QN9_WHEAT	Q75qn9 triticum ae
23	87	28.4	305	2	Q9CX86_MOUSE	Q9cx86 mus musculu
24	86.5	28.3	631	2	Q65AS3_9HEMI	Q65as3 nilaparvata
25	86	28.1	167	2	Q651J0_ORYSA	Q651j0 oryza sativ
26	85.5	27.9	582	2	Q5GQB4_9CAUD	Q5gqb4 bacteriopha
27	85	27.8	231	1	NOLA1_MOUSE	Q9cy66 mus musculu
28	85	27.8	307	2	Q7S7R5_NEUCR	Q7s7r5 neurospora
29	85	27.8	1038	2	Q5Z5B2_ORYSA	Q5z5b2 oryza sativ
30	84.5	27.6	249	2	Q7XGH3_ORYSA	Q7xgh3 oryza sativ
31	84.5	27.6	249	2	Q8S748_ORYSA	Q8s748 oryza sativ
32	84.5	27.6	550	2	O17145_LUCCU	O17145 lucilia cup
33	84	27.5	1610	2	Q92KQ8_RHIME	Q92kq8 rhizobium m
34	83.5	27.3	197	2	Q84UR8_ORYSA	Q84ur8 oryza sativ
35	83.5	27.3	243	2	Q8N7I6_HUMAN	Q8n7i6 homo sapien
36	83.5	27.3	540	2	Q7NL21_GLOVI	Q7nl21 gloeobacter
37	83.5	27.3	970	2	Q9CA47_ARATH	Q9ca47 arabidopsis
38	83.5	27.3	1276	2	Q9SSF7_ARATH	Q9ssf7 arabidopsis
39	83	27.1	89	2	Q6ZAB1_ORYSA	Q6zab1 oryza sativ
40	83	27.1	98	2	Q867H1_9ECHO	Q867h1 echinometra
41	83	27.1	98	2	Q867S5_9ECHO	Q867s5 echinometra
42	83	27.1	98	2	Q86D41_9ECHO	Q86d41 echinometra
43	83	27.1	98	2	Q86D48_9ECHO	Q86d48 echinometra
44	83	27.1	98	2	Q86D51_9ECHO	Q86d51 echinometra
45	83	27.1	98	2	Q86D52_9ECHO	Q86d52 echinometra

ALIGNMENTS

RESULT 1
Q4W7T9_9CRUS
ID Q4W7T9_9CRUS PRELIMINARY; PRT; 775 AA.
AC Q4W7T9;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE VASA RNA helicase.
GN Name=vasa;
OS Daphnia magna.
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Branchiopoda; Diplostraca;
OC Cladocera; Anomopoda; Daphniidae; Daphnia.
OX NCBI_TaxID=35525;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sagawa K., Yamagata H., Shiga Y.;
RT "Exploring embryonic germ line development in the water flea, Daphnia
RT magna, by zinc-finger-containing VASA as a marker.";
RL Gene Expr. Patterns 5:669-678(2005).
DR EMBL; AB193324; BAD99522.1; -; mRNA.
KW Helicase.
SQ SEQUENCE 775 AA; 82164 MW; E388E608BA098125 CRC64;

Query Match 36.9%; Score 113; DB 2; Length 775;
Best Local Similarity 47.4%; Pred. No. 0.003;
Matches 27; Conservative 5; Mismatches 15; Indels 10; Gaps 4;

```
Qy          1 GGTYSCH----FGPLTWVCKPQGGGGGGGGTYSCH----FGPLTWVCKPQGGGGGGG 49
              ||: :||    |  :  | || ||||| |  :||    |  :  | ||||||||||
Db        162 GGSRACHKCGEEGHFSREC-PQAGGGGGSGPRTCHKCGEEGHFSREC-PQGGGGGGG 216
```

RESULT 2
Q4R1V0_9CRUS
ID Q4R1V0_9CRUS PRELIMINARY; PRT; 779 AA.
AC Q4R1V0;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE VASA RNA helicase.
GN Name=vasa;
OS Daphnia magna.
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Branchiopoda; Diplostraca;
OC Cladocera; Anomopoda; Daphniidae; Daphnia.
OX NCBI_TaxID=35525;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sagawa K., Yamagata H., Shiga Y.;
RT "Exploring embryonic germ line development in the water flea, Daphnia
RT magna, by zinc-finger-containing VASA as a marker.";
RL Gene Expr. Patterns 5:669-678(2005).
DR EMBL; AB193327; BAE00180.1; -; Genomic_DNA.
KW Helicase.
SQ SEQUENCE 779 AA; 82342 MW; B6C30D45AA8B352F CRC64;

```
  Query Match             36.9%;  Score 113;  DB 2;  Length 779;
  Best Local Similarity   47.4%;  Pred. No. 0.003;
  Matches   27;  Conservative    5;  Mismatches   15;  Indels   10;  Gaps    4;

Qy          1 GGTYSCH----FGPLTWVCKPQGGGGGGGGTYSCH----FGPLTWVCKPQGGGGGGG 49
              ||: :||     |  :  | || ||||| |  :||     |  :  | |||||||||
Db        163 GGSRACHKCGEEGHFSREC-PQAGGGGGSGPRTCHKCGEEGHFSREC-PQGGGGGGG 217


RESULT 3
Q6YUR8_ORYSA
ID   Q6YUR8_ORYSA PRELIMINARY;      PRT;   241 AA.
AC   Q6YUR8;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE   Putative Glycine-rich protein 2.
GN   Name=OSJNBb0088N06.21; Synonyms=OJ1020_C02.12;
OS   Oryza sativa (japonica cultivar-group).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=39947;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Sasaki T., Matsumoto T., Katayose Y.;
RT   "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 2, BAC
RT   clone:OSJNBb0088N06.";
RL   Submitted (OCT-2002) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RA   Sasaki T., Matsumoto T., Yamamoto K.;
RT   "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 2, BAC
RT   clone:OJ1020_C02.";
RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Contains 1 CSD (cold-shock) domain.
DR   EMBL; AP005851; BAD08139.1; -; Genomic_DNA.
DR   EMBL; AP004078; BAD07599.1; -; Genomic_DNA.
DR   Gramene; Q6YUR8; -.
DR   GO; GO:0003677; F:DNA binding; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR011129; CSP.
DR   InterPro; IPR002059; CSP_DNA_bd.
DR   InterPro; IPR002952; Eggshell.
DR   InterPro; IPR012340; OB_NA_bd_sub.
DR   InterPro; IPR001878; Znf_CCHC.
DR   Pfam; PF00313; CSD; 1.
DR   Pfam; PF00098; zf-CCHC; 4.
DR   PRINTS; PR00939; C2HCZNFINGER.
DR   PRINTS; PR00050; COLDSHOCK.
DR   PRINTS; PR01228; EGGSHELL.
DR   ProDom; PD000621; Cold_shock; 1.
DR   SMART; SM00357; CSP; 1.
DR   SMART; SM00343; ZnF_C2HC; 4.
DR   PROSITE; PS00352; COLD_SHOCK; 1.
DR   PROSITE; PS50158; ZF_CCHC; 4.
KW   RNA-binding.
SQ   SEQUENCE   241 AA;  22723 MW;  69E6A187A7B35E03 CRC64;

  Query Match             33.2%;  Score 101.5;  DB 2;  Length 241;
  Best Local Similarity   43.1%;  Pred. No. 0.014;
  Matches   25;  Conservative    3;  Mismatches   21;  Indels    9;  Gaps    3;

Qy          1 GGTYSC-HFGPLTWVCKPQGGGGGGGG-------TYSC-HFGPLTWVCKPQGGGGGGG 49
              || : |   | :   |  ||||||||        |:|   | |   |  |||||||
Db        159 GGCFKCGEMGHMARDCFNSGGGGGGGGGGGGGGGACYNCGETGHLARDCYNGGGGGGG 216


RESULT 4
Q8LPA7_WHEAT
ID   Q8LPA7_WHEAT PRELIMINARY;      PRT;   229 AA.
AC   Q8LPA7;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Cold shock protein-1.
GN   Name=WCSP1;
OS   Triticum aestivum (Wheat).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC   Triticeae; Triticum.
OX   NCBI_TaxID=4565;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22218063; PubMed=12122010; DOI=10.1074/jbc.M205774200;
RA   Karlson D., Nakaminami K., Toyomasu T., Imai R.;
RT   "A cold-regulated nucleic acid-binding protein of winter wheat shares
RT   a domain with bacterial cold shock proteins.";
RL   J. Biol. Chem. 277:35248-35256(2002).
CC   -!- SIMILARITY: Contains 1 CSD (cold-shock) domain.
DR   EMBL; AB066265; BAB78536.2; -; mRNA.
DR   HSSP; P15277; 1MJC.
DR   Gramene; Q8LPA7; -.
DR   GO; GO:0003677; F:DNA binding; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR011129; CSP.
DR   InterPro; IPR002059; CSP_DNA_bd.
DR   InterPro; IPR012340; OB_NA_bd_sub.
DR   InterPro; IPR001878; Znf_CCHC.
DR   Pfam; PF00313; CSD; 1.
DR   Pfam; PF00098; zf-CCHC; 3.
DR   PRINTS; PR00939; C2HCZNFINGER.
DR   PRINTS; PR00050; COLDSHOCK.
DR   ProDom; PD000621; Cold_shock; 1.
DR   SMART; SM00357; CSP; 1.
DR   SMART; SM00343; ZnF_C2HC; 3.
DR   PROSITE; PS00352; COLD_SHOCK; 1.
DR   PROSITE; PS50158; ZF_CCHC; 3.
KW   RNA-binding.
SQ   SEQUENCE   229 AA;  21384 MW;  4CB5C9B6323BD23C CRC64;

  Query Match             31.7%;  Score 97;  DB 2;  Length 229;
  Best Local Similarity   42.2%;  Pred. No. 0.037;
  Matches   27;  Conservative    4;  Mismatches   15;  Indels   18;  Gaps    5;

Qy          2 GTYSC-HFGPLTWVCKPQGGGGGGG-------------GTYSC-HFGPLTWVCKPQGGG 45
              | | |   | ::  | |||||||||             | |   | ::  | |||||
Db        131 GCYKCGEEGHISRDC-PQGGGGGGGYGGGGYGGGGGGGRECYKCGEEGHISRDC-PQGGG 188

Qy         46 GGGG 49
              ||||
Db        189 GGGG 192


RESULT 5
Q75QN8_WHEAT
ID   Q75QN8_WHEAT PRELIMINARY;      PRT;   231 AA.
AC   Q75QN8;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Cold shock domain protein 3.
GN   Name=WCSP3;
OS   Triticum aestivum (Wheat).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC   Triticeae; Triticum.
OX   NCBI_TaxID=4565;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Nakaminami K., Imai R.;
RL   Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Contains 1 CSD (cold-shock) domain.
DR   EMBL; AB161683; BAD08701.1; -; mRNA.
DR   Gramene; Q75QN8; -.
```

```
DR   GO; GO:0003677; F:DNA binding; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR011129; CSP.
DR   InterPro; IPR002059; CSP_DNA_bd.
DR   InterPro; IPR012340; OB_NA_bd_sub.
DR   InterPro; IPR001878; Znf_CCHC.
DR   Pfam; PF00313; CSD; 1.
DR   Pfam; PF00098; zf-CCHC; 3.
DR   PRINTS; PR00939; C2HCZNFINGER.
DR   PRINTS; PR00050; COLDSHOCK.
DR   ProDom; PD000621; Cold_shock; 1.
DR   SMART; SM00357; CSP; 1.
DR   SMART; SM00343; ZnF_C2HC; 3.
DR   PROSITE; PS00352; COLD_SHOCK; 1.
DR   PROSITE; PS50158; ZF_CCHC; 3.
KW   RNA-binding.
SQ   SEQUENCE   231 AA;  21544 MW;  F6E1FE0104CDE2C6 CRC64;

  Query Match             31.2%;  Score 95.5;  DB 2;  Length 231;
  Best Local Similarity   46.0%;  Pred. No. 0.052;
  Matches   23;  Conservative    2;  Mismatches   10;  Indels   15;  Gaps    3;

Qy          1 GGTYSCHFGPLTWVCKPQGGGGGGGGTYSC-HFGPLTWVCKPQGGGGGGG 49
              || |               |||||| | | |    | ::  | |||||||||
Db        125 GGGYG------------GGGGGGRGCYKCGEDGHISRDC-PQGGGGGGG 160


RESULT 6
Q528N5_MAGGR
ID   Q528N5_MAGGR PRELIMINARY;      PRT;   210 AA.
AC   Q528N5;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Predicted protein.
GN   ORFNames=MG05406.4;
OS   Magnaporthe grisea 70-15.
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC   Sordariomycetes incertae sedis; Magnaporthaceae; Magnaporthe.
OX   NCBI_TaxID=242507;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=70-15;
RA   Birren B., Nusbaum C., Abebe A., Abouelleil A., Adekoya E.,
RA   Ait-zahra M., Allen N., Allen T., An P., Anderson M., Anderson S.,
RA   Arachchi H., Armbruster J., Bachantsang P., Baldwin J., Barry A.,
RA   Bayul T., Blitshsteyn B., Bloom T., Blye J., Boguslavskiy L.,
RA   Borowsky M., Boukhgalter B., Brunache A., Butler J., Calixte N.,
RA   Calvo S., Camarata J., Campo K., Chang J., Cheshatsang Y., Citroen M.,
RA   Collymore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
RA   David R., Dawoe T., Degray S., Dodge S., Dooley K., Dorje P.,
RA   Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,
RA   Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,
RA   Fitzgerald M., Foley K., Gage D., Galagan J., Gearin G., Gnerre S.,
RA   Gnirke A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,
RA   Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
RA   Honan T., Horn A., Houde N., Hughes L., Hulme W., Husby E., Iliev I.,
RA   Jaffe D., Jones C., Kamal M., Kamat A., Kamvysselis M., Karlsson E.,
RA   Kells C., Kieu A., Kisner P., Kodira C., Kulbokas E., Labutti K.,
RA   Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
RA   Lindblad-toh K., Liu X., Lokyitsang T., Lokyitsang Y., Lucien O.,
RA   Lui A., Ma L.J., Mabbitt R., Macdonald J., Maclean C., Major J.,
RA   Manning J., Marabella R., Maru K., Matthews C., Mauceli E.,
RA   Mccarthy M., Mcdonough S., Mcghee T., Meldrim J., Meneus L.,
RA   Mesirov J., Mihalev A., Mihova T., Mikkelsen T., Mlenga V., Moru K.,
RA   Mozes J., Mulrain L., Munson G., Naylor J., Newes C., Nguyen C.,
RA   Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
RA   Norbu N., O'donnell P., Okoawo O., O'leary S., Omotosho B.,
RA   O'neill K., Osman S., Parker S., Perrin D., Phunkhang P., Piqani B.,
RA   Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C.,
RA   Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
RA   Rutman M., Schupbach R., Seaman C., Settipalli S., Sharpe T.,
RA   Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sougnez C.,
RA   Spencer B., Stalker J., Stange-thomann N., Stavropoulos S.,
RA   Stetson K., Stone C., Stone S., Stubbs M., Talamas J., Tchuinga P.,
RA   Tenzing P., Tesfaye S., Theodore J., Thoulutsang Y., Topham K.,
RA   Towey S., Tsamla T., Tsomo N., Vallee D., Vassiliev H.,
RA   Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
RA   Wangdi T., Whittaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,
RA   Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembeck L.,
RA   Zimmer A., Zody M., Lander E.;
RT   "The genome sequence of Magnaporthe grisea.";
RL   Submitted (OCT-2003) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=70-15;
RA   Dean R., Mitchell T., Brown D., Pan H., Thon M.;
RL   Submitted (OCT-2003) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=70-15;
RA   Zhu H., Blackmon B.;
RL   Submitted (OCT-2003) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; AACU01000505; EAA54614.1; -; Genomic_DNA.
SQ   SEQUENCE   210 AA;  21723 MW;  CE4821AC5E07E149 CRC64;

  Query Match             30.7%;  Score 94;  DB 2;  Length 210;
  Best Local Similarity   45.3%;  Pred. No. 0.067;
  Matches   24;  Conservative    3;  Mismatches   20;  Indels    6;  Gaps    2;

Qy          2 GTYSCHFGPLTWVC-----KPQGGGGGGGGTYSCHFGPLTWVCKPQGGGGGGG 49
              || :|| | |||        : |||| |||    |  |     |  :   |||||||
Db        134 GT-NCHPGGLTWDFAIEDRRGQGGGNGGGRATSSAFPGSGWGLRGGIGGGGGG 185


RESULT 7
Q9GNP2_CIOSA
ID   Q9GNP2_CIOSA PRELIMINARY;      PRT;   688 AA.
AC   Q9GNP2;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Vasa homolog.
GN   Name=CsDEAD1a(CsVHa);
OS   Ciona savignyi.
OC   Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
OC   Phlebobranchia; Cionidae; Ciona.
OX   NCBI_TaxID=51511;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Ovary;
RX   MEDLINE=20130953; PubMed=10664149; DOI=10.1007/s004270050012;
RA   Fujimura M., Takamura K.;
RT   "Characterization of an ascidian DEAD-box gene, Ci-DEAD1: specific
RT   expression in the germ cells and its mRNA localization in the
RT   posterior-most blastomeres in early embryos.";
RL   Dev. Genes Evol. 210:64-72(2000).
DR   EMBL; AB047802; BAB12216.1; -; mRNA.
DR   HSSP; Q58083; 1HV8.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR   GO; GO:0003676; F:nucleic acid binding; IEA.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR011545; DEAD/DEAH_N.
DR   InterPro; IPR000629; DEAD_box.
DR   InterPro; IPR001650; Helicase_C.
DR   InterPro; IPR001878; Znf_CCHC.
DR   Pfam; PF00270; DEAD; 1.
DR   Pfam; PF00271; Helicase_C; 1.
DR   Pfam; PF00098; zf-CCHC; 3.
DR   PRINTS; PR00939; C2HCZNFINGER.
DR   SMART; SM00487; DEXDc; 1.
DR   SMART; SM00490; HELICc; 1.
```

```
DR   SMART; SM00343; ZnF_C2HC; 3.
DR   PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
DR   PROSITE; PS50158; ZF_CCHC; 3.
SQ   SEQUENCE   688 AA;  73744 MW;  7EE70CFE04A681E9 CRC64;

  Query Match             29.9%;  Score 91.5;  DB 2;  Length 688;
  Best Local Similarity   47.1%;  Pred. No. 0.38;
  Matches   24;  Conservative    6;  Mismatches   16;  Indels    5;  Gaps    5;

Qy          2 GTYSC-HFGPLTWVCKPQGGGGG-GGGTYSC-HFGPLTWVCKPQGGGGGGG 49
              | : |   | ::  | ||||||| | | : |   | ::  | ||||||| |
Db        107 GCFKCGEEGHMSREC-PQGGGGGRGSGCFKCGEEGHMSREC-PQGGGGGRG 155


RESULT 8
Q9GNP1_CIOSA
ID   Q9GNP1_CIOSA PRELIMINARY;      PRT;   770 AA.
AC   Q9GNP1;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Vasa homolog.
GN   Name=CsDEAD1b(CsVHb);
OS   Ciona savignyi.
OC   Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
OC   Phlebobranchia; Cionidae; Ciona.
OX   NCBI_TaxID=51511;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Ovary;
RX   MEDLINE=20130953; PubMed=10664149; DOI=10.1007/s004270050012;
RA   Fujimura M., Takamura K.;
RT   "Characterization of an ascidian DEAD-box gene, Ci-DEAD1: specific
RT   expression in the germ cells and its mRNA localization in the
RT   posterior-most blastomeres in early embryos.";
RL   Dev. Genes Evol. 210:64-72(2000).
DR   EMBL; AB047803; BAB12217.1; -; mRNA.
DR   HSSP; Q58083; 1HV8.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR   GO; GO:0003676; F:nucleic acid binding; IEA.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR011545; DEAD/DEAH_N.
DR   InterPro; IPR000629; DEAD_box.
DR   InterPro; IPR001650; Helicase_C.
DR   InterPro; IPR001878; Znf_CCHC.
DR   Pfam; PF00270; DEAD; 1.
DR   Pfam; PF00271; Helicase_C; 1.
DR   Pfam; PF00098; zf-CCHC; 6.
DR   PRINTS; PR00939; C2HCZNFINGER.
DR   SMART; SM00487; DEXDc; 1.
DR   SMART; SM00490; HELICc; 1.
DR   SMART; SM00343; ZnF_C2HC; 6.
DR   PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
DR   PROSITE; PS50158; ZF_CCHC; 6.
SQ   SEQUENCE   770 AA;  82032 MW;  5C6D2A2D8C9CDD58 CRC64;

  Query Match             29.9%;  Score 91.5;  DB 2;  Length 770;
  Best Local Similarity   47.1%;  Pred. No. 0.42;
  Matches   24;  Conservative    6;  Mismatches   16;  Indels    5;  Gaps    5;

Qy          2 GTYSC-HFGPLTWVCKPQGGGGG-GGGTYSC-HFGPLTWVCKPQGGGGGGG 49
              | : |   | ::  | ||||||| | | : |   | ::  | ||||||| |
Db        189 GCFKCGEEGHMSREC-PQGGGGGRGSGCFKCGEEGHMSREC-PQGGGGGRG 237


RESULT 9
Q9GV13_HYDMA
ID   Q9GV13_HYDMA PRELIMINARY;      PRT;   797 AA.
AC   Q9GV13;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Vasa-related protein CnVAS1.
GN   Name=Cnvas1;
OS   Hydra magnipapillata (Hydra).
OC   Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Anthomedusae;
OC   Hydridae; Hydra.
OX   NCBI_TaxID=6085;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=21359115; PubMed=11466525; DOI=10.1007/s004270100156;
RA   Mochizuki K., Nishimiya-Fujisawa C., Fujisawa T.;
RT   "Universal occurrence of the vasa-related genes among metazoans and
RT   their germline expression in Hydra.";
RL   Dev. Genes Evol. 211:299-308(2001).
DR   EMBL; AB047382; BAB13307.1; -; mRNA.
DR   HSSP; Q58083; 1HV8.
DR   GO; GO:0005524; F:ATP binding; IEA.
DR   GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR   GO; GO:0003676; F:nucleic acid binding; IEA.
DR   InterPro; IPR001410; DEAD.
DR   InterPro; IPR011545; DEAD/DEAH_N.
DR   InterPro; IPR000629; DEAD_box.
DR   InterPro; IPR001650; Helicase_C.
DR   InterPro; IPR001878; Znf_CCHC.
DR   Pfam; PF00270; DEAD; 1.
DR   Pfam; PF00271; Helicase_C; 1.
DR   Pfam; PF00098; zf-CCHC; 7.
DR   PRINTS; PR00939; C2HCZNFINGER.
DR   SMART; SM00487; DEXDc; 1.
DR   SMART; SM00490; HELICc; 1.
DR   SMART; SM00343; ZnF_C2HC; 7.
DR   PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
DR   PROSITE; PS50158; ZF_CCHC; 7.
SQ   SEQUENCE   797 AA;  83935 MW;  847D939DA1AF622C CRC64;

  Query Match             29.9%;  Score 91.5;  DB 2;  Length 797;
  Best Local Similarity   32.5%;  Pred. No. 0.44;
  Matches   25;  Conservative    7;  Mismatches   16;  Indels   29;  Gaps    4;

Qy          1 GGTYSCH----FGPLTWVCKPQGGGG-------------------GGGGTYSCH----F 32
              ||  :||     | ::  |   ||||                     ||||: :||
Db        115 GGGRACHKCGKEGHMSRECPDGGGGGRACFKCKQEGHMSKDCPQGSGGGGSRTCHKCGKE 174

Qy         33 GPLTWVCKPQGGGGGGG 49
              | ::  | | | |||||
Db        175 GHMSREC-PDGSGGGGG 190


RESULT 10
Q7XJR8_ARATH
ID   Q7XJR8_ARATH PRELIMINARY;      PRT;   299 AA.
AC   Q7XJR8;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Putative glycine-rich, zinc-finger DNA-binding protein.
GN   Name=At2g17870;
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC   rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=20083487; PubMed=10617197; DOI=10.1038/45471;
RA   Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA   Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA   Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H.L.,
RA   Moffat K.S., Cronin L.A., Shen M., Pai G., Van Aken S., Umayam L.,
RA   Tallon L.J., Gill J.E., Adams M.D., Carrera A.J., Creasy T.H.,
RA   Goodman H.M., Somerville C.R., Copenhaver G.P., Preuss D.,
RA   Nierman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,
RA   Venter J.C.;
RT   "Sequence and analysis of chromosome 2 of the plant Arabidopsis
```

```
RT   thaliana.";
RL   Nature 402:761-768(1999).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RA   Lin X.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC003952; AAD03571.1; -; Genomic_DNA.
DR   HSSP; O54310; 1G6P.
DR   GO; GO:0003677; F:DNA binding; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR002059; Cold_shock.
DR   InterPro; IPR011129; CSP.
DR   InterPro; IPR001878; Znf_CCHC.
DR   Pfam; PF00313; CSD; 1.
DR   Pfam; PF00098; zf-CCHC; 7.
DR   PRINTS; PR00939; C2HCZNFINGER.
DR   PRINTS; PR00050; COLDSHOCK.
DR   ProDom; PD000621; Cold_shock; 1.
DR   SMART; SM00357; CSP; 1.
DR   SMART; SM00343; ZnF_C2HC; 7.
DR   PROSITE; PS00352; COLD_SHOCK; 1.
DR   PROSITE; PS50158; ZF_CCHC; 7.
KW   DNA-binding; Zinc-finger.
SQ   SEQUENCE   299 AA;  29362 MW;  3B1624E567CAE73F CRC64;

  Query Match             29.7%;  Score 91;  DB 2;  Length 299;
  Best Local Similarity   44.8%;  Pred. No. 0.19;
  Matches   26;  Conservative    3;  Mismatches   15;  Indels   14;  Gaps    6;

Qy          2 GTYSC----HFGPLTWVCKPQGGG--GGGGGT-YSC-HFGPLTWVCK---PQGGGGGG 48
              | | |    ||      |:  |||  |||| | |:|   | :  ||    | ||||||
Db        195 GCYMCGGVGHFAR---DCRQNGGGNVGGGGSTCYTCGGVGHIAKVCTSKIPSGGGGGG 249


RESULT 11
Q94C69_ARATH
ID   Q94C69_ARATH PRELIMINARY;      PRT;   301 AA.
AC   Q94C69;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE   Putative glycine-rich, zinc-finger DNA-binding protein.
GN   Name=At2g17870;
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
RA   Goldsmith A.D., Lee J.M., Quach H.L., Toriumi M., Yu G., Bowser L.,
RA   Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA   Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA   Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,
RA   Seki M., Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
RA   Theologis A.;
RL   Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RA   Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Etgu P., Lee J.M.,
RA   Toriumi M., Yu G., Brooks S., Chao Q., Chen H., Karlin-Neumann G.,
RA   Kim C., Lam B., Miranda M., Nguyen M., Palm C.J., Shinn P.,
RA   Southwick A., Davis R.W., Ecker J.R., Theologis A.;
RL   Submitted (NOV-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY035133; AAK59638.1; -; mRNA.
DR   EMBL; AY062985; AAL34159.1; -; mRNA.
DR   PIR; T00837; T00837.
DR   HSSP; O54310; 1G6P.
DR   GO; GO:0003677; F:DNA binding; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR002059; Cold_shock.
DR   InterPro; IPR011129; CSP.
DR   InterPro; IPR001878; Znf_CCHC.
DR   Pfam; PF00313; CSD; 1.
DR   Pfam; PF00098; zf-CCHC; 7.
DR   PRINTS; PR00939; C2HCZNFINGER.
DR   PRINTS; PR00050; COLDSHOCK.
DR   ProDom; PD000621; Cold_shock; 1.
DR   SMART; SM00357; CSP; 1.
DR   SMART; SM00343; ZnF_C2HC; 7.
DR   PROSITE; PS00352; COLD_SHOCK; 1.
DR   PROSITE; PS50158; ZF_CCHC; 7.
KW   DNA-binding; Zinc-finger.
SQ   SEQUENCE   301 AA;  29564 MW;  28F6A32F4C48CFBF CRC64;

  Query Match             29.7%;  Score 91;  DB 2;  Length 301;
  Best Local Similarity   44.8%;  Pred. No. 0.19;
  Matches   26;  Conservative    3;  Mismatches   15;  Indels   14;  Gaps    6;

Qy          2 GTYSC----HFGPLTWVCKPQGGG--GGGGGT-YSC-HFGPLTWVCK---PQGGGGGG 48
              | | |    ||      |:  |||  |||| | |:|   | :  ||    | ||||||
Db        197 GCYMCGGVGHFAR---DCRQNGGGNVGGGGSTCYTCGGVGHIAKVCTSKIPSGGGGGG 251


RESULT 12
Q6YUX6_ORYSA
ID   Q6YUX6_ORYSA PRELIMINARY;      PRT;   410 AA.
AC   Q6YUX6;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE   Putative TCP-domain protein.
GN   Name=OSJNBa0078N11.17; Synonyms=OSJNBb0024K03.29;
OS   Oryza sativa (japonica cultivar-group).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=39947;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Sasaki T., Matsumoto T., Katayose Y.;
RL   Submitted (OCT-2002) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RA   Sasaki T., Matsumoto T., Katayose Y.;
RT   "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 2, BAC
RT   clone:OSJNBb0024K03.";
RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AP005848; BAD16444.1; -; Genomic_DNA.
DR   EMBL; AP005733; BAD16334.1; -; Genomic_DNA.
DR   Gramene; Q6YUX6; -.
DR   InterPro; IPR005333; TCP.
DR   Pfam; PF03634; TCP; 1.
SQ   SEQUENCE   410 AA;  41249 MW;  1073C57678220D98 CRC64;

  Query Match             29.2%;  Score 89.5;  DB 2;  Length 410;
  Best Local Similarity   47.7%;  Pred. No. 0.36;
  Matches   21;  Conservative    0;  Mismatches   12;  Indels   11;  Gaps    1;

Qy         17 PQGGGGGGGGTYSCHFGPL----------TWVCKPQGGGGGGG 49
              |||||||||    | | |          |     ||||||||
Db        342 PVGGGGGGGGGGEGHMGILAALNAYRTQAATDAAGQQGGGGGGG 385


RESULT 13
Q7YXK1_ASCSU
ID   Q7YXK1_ASCSU PRELIMINARY;      PRT;   472 AA.
AC   Q7YXK1;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   MFP3.
OS   Ascaris suum (Pig roundworm) (Ascaris lumbricoides).
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea;
OC   Ascarididae; Ascaris.
```

```
OX   NCBI_TaxID=6253;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Testis;
RX   MEDLINE=23001057; PubMed=14565983; DOI=10.1091/mbc.E03-04-0246;
RA   Buttery S.M., Ekman G.C., Seavy M., Stewart M., Roberts T.M.;
RT   "Dissection of the Ascaris sperm motility machinery identifies key
RT   proteins involved in major sperm protein-based amoeboid locomotion.";
RL   Mol. Biol. Cell 14:5082-5088(2003).
DR   EMBL; AY326288; AAP94887.1; -; mRNA.
SQ   SEQUENCE   472 AA;  43258 MW;  E594625F523A73EC CRC64;

  Query Match             29.2%;  Score 89.5;  DB 2;  Length 472;
  Best Local Similarity   42.9%;  Pred. No. 0.42;
  Matches   21;  Conservative    2;  Mismatches    9;  Indels   17;  Gaps    2;

Qy          1 GGTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGGGGGGG 49
              |||  | :||          ||    || ||:||          |||||||
Db        403 GGTTSAYFGV---------GGSSAPGTTSCYFG--------AGGGGGGG 434


RESULT 14
Q9VD49_DROME
ID   Q9VD49_DROME PRELIMINARY;      PRT;   117 AA.
AC   Q9VD49;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   CG5778-PA (GH13168p).
GN   ORFNames=CG5778, CG5778;
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22426065; PubMed=12537568;
RA   Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA   Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA   George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,
RA   Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA   Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA   Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT   "Finishing a whole-genome shotgun: release 3 of the Drosophila
RT   melanogaster euchromatic genome sequence.";
RL   Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22426070; PubMed=12537573;
RA   Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J.W., Svirskas R.,
RA   Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA   Ashburner M., Celniker S.E.;
RT   "The transposable elements of the Drosophila melanogaster euchromatin:
RT   a genomics perspective.";
RL   Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).
RN   [4]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22426069; PubMed=12537572;
RA   Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA   Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA   Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA   Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA   Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA   Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA   Lewis S.E.;
RT   "Annotation of the Drosophila melanogaster euchromatic genome: a
RT   systematic review.";
RL   Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN   [5]
RP   NUCLEOTIDE SEQUENCE.
RG   Berkeley Drosophila Genome Project;
RA   Celniker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,
RA   Hoskins R., Stapleton M., Pacleb J., Park S., Svirskas R., Smith E.,
RA   Yu C., Rubin G.;
RT   "Drosophila melanogaster release 4 sequence.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   NUCLEOTIDE SEQUENCE.
RG   FlyBase;
RL   Submitted (MAR-2005) to the EMBL/GenBank/DDBJ databases.
RN   [7]
RP   NUCLEOTIDE SEQUENCE;
RC   STRAIN=Berkeley;
RA   Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA   Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA   George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA   Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA   Patel S., Phouanenavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA   Celniker S.;
RL   Submitted (JUN-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE003737; AAF55954.1; -; Genomic_DNA.
DR   EMBL; AY118781; AAM50641.1; -; mRNA.
DR   Ensembl; CG5778; Drosophila melanogaster.
DR   FlyBase; FBgn0038930; CG5778.
SQ   SEQUENCE   117 AA;  10632 MW;  3AE2D79CA8924A96 CRC64;

  Query Match             29.1%;  Score 89;  DB 2;  Length 117;
  Best Local Similarity   44.2%;  Pred. No. 0.12;
  Matches   23;  Conservative    2;  Mismatches    7;  Indels   20;  Gaps    3;

Qy          1 GGTYSCHFGPLTWVCKPQGGGGGGGG---TYSCHFGPLTWVCKPQGGGGGGG 49
              || |:             ||||||||    || :|||        |||||||
Db         62 GGYYN------------GGGGGGGGRRPVYSGNFGP----GYGNGGGGGGG 96
```

```
RESULT 15
NKX23_MOUSE
ID   NKX23_MOUSE    STANDARD;      PRT;   362 AA.
AC   P97334; Q9QZ60; Q9WV67;
DT   15-JUL-1998 (Rel. 36, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-MAY-2005 (Rel. 47, Last annotation update)
DE   Homeobox protein Nkx-2.3 (Homeobox protein NK-2 homolog C) (Nkx2-C)
DE   (Homeobox protein NK-2 homolog 3).
GN   Name=Nkx2-3; Synonyms=Nkx-2.3, Nkx2c;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC   Muroidea; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=129; TISSUE=Liver;
RX   MEDLINE=97287401; PubMed=9142493;
RX   DOI=10.1002/(SICI)1097-0177(199705)209:1<29::AID-AJA3>3.3.CO;2-X;
RA   Pabst O., Schneider A., Brand T., Arnold H.-H.;
RT   "The mouse Nkx2-3 homeodomain gene is expressed in gut mesenchyme
RT   during pre- and postnatal mouse development.";
RL   Dev. Dyn. 209:29-35(1997).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=129;
RA   Wang C.-C., Biben C., Robb R., Tarlinton D., Nassir F., Davidson N.O.,
RA   Harvey R.P.;
RT   "Homeodomain factor Nkx2-3 is required for normal development of the
RT   gut-associated lymphoid tissue and the spleen.";
RL   Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=BALB/c;
RX   MEDLINE=20387131; PubMed=10926756; DOI=10.1006/dbio.2000.9749;
RA   Wang C.-C., Biben C., Robb L., Nassir F., Barnett L., Davidson N.O.,
RA   Koentgen F., Tarlinton D., Harvey R.P.;
RT   "Homeodomain factor Nkx2-3 controls regional expression of leukocyte
RT   homing coreceptor MAdCAM-1 in specialized endothelial cells of the
RT   viscera.";
RL   Dev. Biol. 224:152-167(2000).
RN   [4]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC   STRAIN=C57BL/6; TISSUE=Brain;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [5]
RP   FUNCTION.
RX   MEDLINE=20253077; PubMed=10790368; DOI=10.1093/emboj/19.9.2015;
RA   Pabst O., Foerster R., Lipp M., Engel H., Arnold H.-H.;
RT   "NKX2.3 is required for MAdCAM-1 expression and homing of lymphocytes
RT   in spleen and mucosa-associated lymphoid tissue.";
RL   EMBO J. 19:2015-2023(2000).
RN   [6]
RP   TISSUE DISTRIBUTION.
RX   MEDLINE=22136510; PubMed=12141427;
RA   Biben C., Wang C.-C., Harvey R.P.;
RT   "NK-2 class homeobox genes and pharyngeal/oral patterning: Nkx2-3 is
RT   required for salivary gland and tooth morphogenesis.";
RL   Int. J. Dev. Biol. 46:415-422(2002).
CC   -!- FUNCTION: Transcriptional regulator essential for normal
CC       development and functions of the small intestine and spleen.
CC       Activates directly MADCAM1 expression. Required for homing of
CC       lymphocytes in spleen and mucosa-associated lymphoid tissue. May
CC       have a role during pharyngeal organogenesis.
CC   -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC   -!- TISSUE SPECIFICITY: Expressed in spleen and intestine. Also
CC       expressed in salivary gland and tongue, which are derivate of the
CC       pharyngeal region.
CC   -!- DEVELOPMENTAL STAGE: Expressed in gut mesenchyme during pre- and
CC       postnatal development. Expressed as well in the pharyngeal floor
CC       and pouches, and in the oral and brancial arch ectoderm.
CC       Expression persisted in the developing thyroid until birth, in
CC       mucous forming cells of salivary glands and in odontogenic
CC       epithelium of the mandibule.
CC   -!- SIMILARITY: Belongs to the NK-2 homeobox family.
CC   -!- SIMILARITY: Contains 1 homeobox DNA-binding domain.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; Y11117; CAA72002.1; -; mRNA.
DR   EMBL; AF202036; AAF08008.1; -; mRNA.
DR   EMBL; AF155583; AAD38415.1; -; Genomic_DNA.
DR   EMBL; BC072614; AAH72614.1; -; mRNA.
DR   HSSP; P23441; 1FTT.
DR   SMR; P97334; 145-210.
DR   TRANSFAC; T04325; -.
DR   Ensembl; ENSMUSG00000044220; Mus musculus.
DR   MGI; MGI:97348; Nkx2-3.
DR   GO; GO:0030183; P:B cell differentiation; IMP.
DR   GO; GO:0006955; P:immune response; IMP.
DR   InterPro; IPR001356; Homeobox.
DR   InterPro; IPR012287; Homeodomain-rel.
DR   InterPro; IPR000047; HTH_lambrepressr.
DR   Pfam; PF00046; Homeobox; 1.
DR   PRINTS; PR00024; HOMEOBOX.
DR   PRINTS; PR00031; HTHREPRESSR.
DR   ProDom; PD000010; Homeobox; 1.
DR   SMART; SM00389; HOX; 1.
DR   PROSITE; PS00027; HOMEOBOX_1; 1.
DR   PROSITE; PS50071; HOMEOBOX_2; 1.
KW   Developmental protein; DNA-binding; Homeobox; Nuclear protein;
KW   Transcription; Transcription regulation.
FT   DNA_BIND     145    204       Homeobox.
FT   COMPBIAS      59     67       Poly-Glu.
FT   COMPBIAS     216    220       Poly-Pro.
FT   COMPBIAS     268    283       Poly-Ala.
FT   COMPBIAS     295    302       Poly-Gly.
FT   CONFLICT      44     44       C -> R (in Ref. 1).
FT   CONFLICT      61     61       D -> E (in Ref. 1).
FT   CONFLICT     132    132       K -> R (in Ref. 1).
FT   CONFLICT     248    250       GVG -> RC (in Ref. 1).
FT   CONFLICT     261    357       AYGYGNSAAAAAAAAAAAAAAAAAYSGSYGCAYPTGGGGGGG
FT                                 GTASAATTAMQPACSATGGGSFVNVSNLGGFGSGGGAQPLH
FT                                 QGAAAGSACTQGTLQ -> TATGTRPPQPLQPPQPQQQGLQ
FT                                 RQLRLAYRPVAAVVVAARPPRRPPPCNPPAAPPGDLRVTAG
FT                                 IEAFVPLVR (in Ref. 1).
FT   CONFLICT     349    349       S -> A (in Ref. 3).
SQ   SEQUENCE   362 AA;  38090 MW;  48886528EC3E1622 CRC64;

  Query Match             28.9%;  Score 88.5;  DB 1;  Length 362;
```

```
  Best Local Similarity   35.6%;  Pred. No. 0.41;
  Matches   21;  Conservative    3;  Mismatches   16;  Indels   19;  Gaps    2;

Qy          2 GTYSCHFGPLTWVCKPQGGGGGGGGTYSCHFGPLTWVCKPQGGG----------GGGG 49
              |:| | :        | |||||||||  |     :   |   |||          | ||
Db        286 GSYGCAY--------PTGGGGGGGGTASAATTAMQPACSATGGGSFVNVSNLGGFGSGG 336


Search completed: March 31, 2006, 16:35:12
Job time : 186.567 secs
```

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: March 31, 2006, 17:35:23 ; Search time 187 Seconds
(without alignments)
37.594 Million cell updates/sec

Title: US-10-609-217-419
Perfect score: 36
Sequence: 1 XYXXXXGPXTWXXXXX 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	34	94.4	521	4	AAB48788	Aab48788 Acidother
2	33	91.7	13	8	ADL92152	Adl92152 Erythropo
3	33	91.7	17	9	ADY54197	Ady54197 Amino aci
4	33	91.7	20	9	ADY54180	Ady54180 Amino aci
5	33	91.7	80	2	AAW14594	Aaw14594 LcnB bact
6	33	91.7	2444	4	ABB71807	Abb71807 Drosophil
7	32	88.9	17	9	ADY54195	Ady54195 Amino aci
8	32	88.9	17	9	ADY54194	Ady54194 Amino aci
9	32	88.9	17	9	ADY54196	Ady54196 Amino aci
10	32	88.9	19	2	AAY26410	Aay26410 Erythropo
11	32	88.9	19	2	AAY13631	Aay13631 Erythropo
12	32	88.9	19	2	AAW26967	Aaw26967 Monomer s
13	32	88.9	19	3	AAB17931	Aab17931 EPO-mimet
14	32	88.9	19	3	AAB17319	Aab17319 EPO-mimet
15	32	88.9	19	3	AAB13505	Aab13505 Erythropo
16	32	88.9	19	5	ABB72828	Abb72828 Erythropo
17	32	88.9	19	5	ABB72844	Abb72844 Erythropo
18	32	88.9	19	5	AAU74460	Aau74460 Human ery
19	32	88.9	19	7	ADJ72583	Adj72583 EPO mimet
20	32	88.9	19	7	ADJ72567	Adj72567 EPO mimet
21	32	88.9	19	8	ADJ52219	Adj52219 CH1 delet
22	32	88.9	19	8	ADJ52203	Adj52203 CH1 delet
23	32	88.9	19	8	ADJ51165	Adj51165 CH1 delet
24	32	88.9	19	8	ADJ51181	Adj51181 CH1 delet
25	32	88.9	19	9	ADZ44412	Adz44412 Erythropo
26	32	88.9	20	2	AAY13696	Aay13696 Erythropo
27	32	88.9	20	2	AAY13650	Aay13650 Erythropo
28	32	88.9	20	2	AAY13728	Aay13728 Erythropo
29	32	88.9	20	2	AAY13687	Aay13687 Erythropo
30	32	88.9	20	2	AAY13661	Aay13661 Erythropo
31	32	88.9	20	2	AAW27001	Aaw27001 Monomer s
32	32	88.9	20	2	AAW27010	Aaw27010 Monomer s
33	32	88.9	20	5	AAU74489	Aau74489 Human ery
34	32	88.9	20	5	AAU74480	Aau74480 Human ery
35	32	88.9	20	9	ADY54174	Ady54174 Amino aci
36	32	88.9	20	9	ADY54157	Ady54157 Amino aci
37	32	88.9	20	9	ADY54181	Ady54181 Amino aci
38	32	88.9	21	9	ADU91978	Adu91978 EPO-R ago
39	32	88.9	22	2	AAY13709	Aay13709 Erythropo
40	32	88.9	22	2	AAY26491	Aay26491 Erythropo
41	32	88.9	22	2	AAY26355	Aay26355 Erythropo
42	32	88.9	22	2	AAW27023	Aaw27023 Monomer s
43	32	88.9	22	9	ADU91911	Adu91911 EPO-R ago
44	32	88.9	23	9	ADU91986	Adu91986 EPO-R ago
45	32	88.9	23	9	ADU91945	Adu91945 EPO-R ago

ALIGNMENTS

RESULT 1
AAB48788
ID AAB48788 standard; protein; 521 AA.
XX
AC AAB48788;
XX
DT 09-MAR-2001 (first entry)
XX
DE Acidothermus cellulolyticus E1 endoglucanase mutant, Y245G.
XX
KW E1 endoglucanase; glycosyl hydrolase; insoluble substrate;
KW cellulose hydrolysis; ethanol production; fermentation; mutant; mutein.
XX
OS Acidothermus cellulolyticus.
OS Synthetic.
XX
PN WO200070031-A1.
XX
PD 23-NOV-2000.
XX
PF 19-MAY-2000; 2000WO-US013971.
XX
PR 19-MAY-1999; 99US-0134925P.
XX
PA (MIDE) MIDWEST RES INST.
XX
PI Himmel ME, Adney WS, Baker JO, Vinzant TB, Thomas SR, Sakon J;
PI Decker SR;
XX
DR WPI; 2001-061226/07.
XX
PT Preparation of glycosyl hydrolase with an increased catalytic activity on
PT insoluble substrate.
XX
PS Claim 5; Page 27-29; 30pp; English.
XX
CC The invention relates to a method for making glycosyl hydrolase mutants
CC with increased catalytic activity with either insoluble or soluble
CC cellulose substrates relative to the wild-type enzyme. The method for
CC making a glycosyl hydrolase with increased insoluble substrate catalytic
CC activity comprises replacing an active site-associated amino glycosyl-
CC stabilising amino acid with an amino acid that does not strongly bind a
CC disaccharide product in the active site. Conversely, the method for
CC making a glycosyl hydrolase with increased soluble substrate catalytic
CC activity comprises replacing a hydrophobic substrate-binding amino acid
CC with a positively charged residue. The invention also discloses mutants
CC of Acidothermus cellulolyticus E1 endoglucanase (AAB48786-B48788)

CC produced according to the method of the invention. The Y245G mutant
CC (AAB48788) has improved activity with insoluble substrates, and the W42R
CC (AAB48786) and Y82R (AAB48787) mutants have improved activity with
CC soluble substrates. The invention also encompasses DNA encoding these
CC mutants. The glycosyl hydrolases of the invention are used as catalysts
CC for cellulose hydrolysis to produce sugars that can be fermented to
CC produce fuels such as ethanol. The present sequence represents the
CC Acidothermus cellulolyticus E1 endoglucanase Y245G mutant
XX
SQ Sequence 521 AA;

Query Match 94.4%; Score 34; DB 4; Length 521;
Best Local Similarity 50.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 YXXXXGPXTW 11
| || ||
Db 240 YATSVGPQTW 249

RESULT 2
ADL92152
ID ADL92152 standard; peptide; 13 AA.
XX
AC ADL92152;
XX
DT 20-MAY-2004 (first entry)
XX
DE Erythropoietin peptide fragment 2.
XX
KW harvesting; recombinant; host cell; N-terminal leader peptide;
KW pre-peptide; lantibiotic; post-translational modification;
KW pharmaceuticals; vaccine; immunogenic.
XX
OS Unidentified.
XX
PN WO2003099862-A1.
XX
PD 04-DEC-2003.
XX
PF 26-MAY-2003; 2003WO-NL000389.
XX
PR 24-MAY-2002; 2002EP-00077060.
PR 07-FEB-2003; 2003US-00360101.
XX
PA (NANO-) APPLIED NANOSYSTEMS BV.
XX
PI Moll GN, Leenhouts CJ, Kuipers OP, Driessen AJM;
XX
DR WPI; 2004-042770/04.
XX
PT Harvesting a desired polypeptide produced by a recombinant host cell, for
PT producing pharmaceuticals, comprises selecting a recombinant nucleic acid
PT comprising nucleic acid fragments encoding a leader peptide and the
PT polypeptide.
XX
PS Claim 4; Page 68; 109pp; English.
XX
CC The invention relates to a novel method for harvesting a (poly)peptide
CC produced by a recombinant host cell. The novel method involves selecting
CC a cell comprising a first nucleic acid encoding a leader peptide and a
CC second nucleic acid fragment encoding the desired (poly)peptide. The
CC first and second fragments are within the same open reading frame of the
CC first nucleic acid and the leader peptide is functionally equivalent to
CC an N-terminal leader peptide found with the pre-peptide of a lantibiotic.
CC The host cells and nucleic acids are useful for producing, harvesting and
CC post-translational modification of polypeptides. The polypeptides may be
CC used in the production of pharmaceuticals, e.g. as antigen for vaccine or
CC immunogenic composition. This sequence represents a polypeptide relating
CC to the novel method of the invention.
XX
SQ Sequence 13 AA;

Query Match 91.7%; Score 33; DB 8; Length 13;
Best Local Similarity 50.0%; Pred. No. 7.8;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 YXXXXGPXTW 11
| || ||
Db 1 YASHFGPLTW 10

RESULT 3
ADY54197
ID ADY54197 standard; peptide; 17 AA.
XX
AC ADY54197;
XX
DT 19-MAY-2005 (first entry)
XX
DE Amino acid sequence of mutated EMP-1 #4.
XX
KW cytostatic; anti-HIV; hypotensive; neuroprotective; cardiovascular-Gen.;
KW nootropic; hepatotropic; virucide; antiinflammatory; immunosuppressive;
KW antiallergic; antimicrobial; neuroleptic; gynecological; anorectic;
KW antiarteriosclerotic; gastrointestinal-Gen.; endocrine-Gen; neoplasm;
KW hematological disease; erythropoietin peptide mimetic; EPM;
KW EPO mimetic peptide-1; EMP-1; multiple sclerosis; brain tumor; cancer;
KW hepatitis; anemia; pregnancy; menstrual disorder; rheumatoid arthritis;
KW AIDS; viral disease; metabolic disease; autoimmune disease;
KW inflammatory disease; allergy; microbial infection;
KW cardiovascular disease; genetic disease; neurodegenerative disease;
KW hematopoietic cell disorder; endocrine disorder;
KW gastrointestinal disease; hypertension; arterial sclerosis.
XX
OS Synthetic.
XX
PN WO2005021579-A2.
XX
PD 10-MAR-2005.
XX
PF 30-AUG-2004; 2004WO-US027949.
XX
PR 28-AUG-2003; 2003WO-US026818.
PR 10-MAR-2004; 2004US-0551552P.
XX
PA (BIOR-) BIOREXIS PHARM CORP.
XX
PI Sadeghi H, Turner AJ;
XX
DR WPI; 2005-214540/22.
XX
PT Novel erythropoietin (EPO) peptide mimetic, having first modification of
PT cysteine residue of EPO mimetic peptides (EMP)-1, to reduces disulfide
PT bond formation, and second modification such that peptide exhibits EMP-1
PT activity.
XX
PS Example 2; SEQ ID NO 51; 158pp; English.
XX
CC The specification describes an erythropoietin (EPO) peptide mimetic
CC (EPM), comprising a modification of at least one cysteine residue of EPO
CC mimetic peptide (EMP)-1 that substantially reduces disulfide bond
CC formation, and a second modification such that the peptide exhibits EMP-1
CC activity. The first modification comprises the deletion or substitution
CC of at least one cysteine residue in EMP-1, and the second modification
CC comprises the addition of a linker group that is covalently bonded to the
CC C-terminal amino acid or N-terminal amino acid of EMP-1. EPM peptides of
CC the invention are useful for treating or preventing diseases, such as
CC multiple sclerosis, brain tumor, skin cancer, hepatitis B, hepatitis C,
CC anemia, beta-thalassemia, pregnancy or menstrual disorders, rheumatoid
CC arthritis, AIDS, cancer, viral disease, metabolic disease, obesity,
CC autoimmune disease, inflammatory disease, allergy, graft-versus-host
CC disease, systemic microbial infection, cardiovascular disease, psychosis,
CC genetic diseases, neurodegenerative diseases, disorders of hematopoietic
CC cells, diseases of the endocrine system or reproductive systems,
CC gastrointestinal diseases, diabetes, asthma, or HIV infections,

CC hypertension, hypercholesterolemia, arterial sclerosis, arthritis or
CC Alzheimer's disease. The present sequence represents a mutated EMP-1,
CC used to produce an EPM of the invention.
XX
SQ Sequence 17 AA;

Query Match 91.7%; Score 33; DB 9; Length 17;
Best Local Similarity 50.0%; Pred. No. 10;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 YXXXXGPXTW 11
| || ||
Db 4 YSGHFGPTTW 13

RESULT 4
ADY54180
ID ADY54180 standard; peptide; 20 AA.
XX
AC ADY54180;
XX
DT 19-MAY-2005 (first entry)
XX
DE Amino acid sequence of a modified EMP-1.
XX
KW cytostatic; anti-HIV; hypotensive; neuroprotective; cardiovascular-Gen.;
KW nootropic; hepatotropic; virucide; antiinflammatory; immunosuppressive;
KW antiallergic; antimicrobial; neuroleptic; gynecological; anorectic;
KW antiarteriosclerotic; gastrointestinal-Gen.; endocrine-Gen; neoplasm;
KW hematological disease; erythropoietin peptide mimetic; EPM;
KW EPO mimetic peptide-1; EMP-1; multiple sclerosis; brain tumor; cancer;
KW hepatitis; anemia; pregnancy; menstrual disorder; rheumatoid arthritis;
KW AIDS; viral disease; metabolic disease; autoimmune disease;
KW inflammatory disease; allergy; microbial infection;
KW cardiovascular disease; genetic disease; neurodegenerative disease;
KW hematopoietic cell disorder; endocrine disorder;
KW gastrointestinal disease; hypertension; arterial sclerosis.
XX
OS Synthetic.
XX
PN WO2005021579-A2.
XX
PD 10-MAR-2005.
XX
PF 30-AUG-2004; 2004WO-US027949.
XX
PR 28-AUG-2003; 2003WO-US026818.
PR 10-MAR-2004; 2004US-0551552P.
XX
PA (BIOR-) BIOREXIS PHARM CORP.
XX
PI Sadeghi H, Turner AJ;
XX
DR WPI; 2005-214540/22.
XX
PT Novel erythropoietin (EPO) peptide mimetic, having first modification of
PT cysteine residue of EPO mimetic peptides (EMP)-1, to reduces disulfide
PT bond formation, and second modification such that peptide exhibits EMP-1
PT activity.
XX
PS Example 1; SEQ ID NO 34; 158pp; English.
XX
CC The specification describes an erythropoietin (EPO) peptide mimetic
CC (EPM), comprising a modification of at least one cysteine residue of EPO
CC mimetic peptide (EMP)-1 that substantially reduces disulfide bond
CC formation, and a second modification such that the peptide exhibits EMP-1
CC activity. The first modification comprises the deletion or substitution
CC of at least one cysteine residue in EMP-1, and the second modification
CC comprises the addition of a linker group that is covalently bonded to the
CC C-terminal amino acid or N-terminal amino acid of EMP-1. EPM peptides of
CC the invention are useful for treating or preventing diseases, such as
CC multiple sclerosis, brain tumor, skin cancer, hepatitis B, hepatitis C,
CC anemia, beta-thalassemia, pregnancy or menstrual disorders, rheumatoid
CC arthritis, AIDS, cancer, viral disease, metabolic disease, obesity,
CC autoimmune disease, inflammatory disease, allergy, graft-versus-host
CC disease, systemic microbial infection, cardiovascular disease, psychosis,
CC genetic diseases, neurodegenerative diseases, disorders of hematopoietic
CC cells, diseases of the endocrine system or reproductive systems,
CC gastrointestinal diseases, diabetes, asthma, or HIV infections,
CC hypertension, hypercholesterolemia, arterial sclerosis, arthritis or
CC Alzheimer's disease. The present sequence represents a modified EMP-1,
CC used to produce an EPM of the invention.
XX
SQ Sequence 20 AA;

Query Match 91.7%; Score 33; DB 9; Length 20;
Best Local Similarity 50.0%; Pred. No. 12;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 YXXXXGPXTW 11
| || ||
Db 4 YSSHFGPLTW 13

RESULT 5
AAW14594
ID AAW14594 standard; protein; 80 AA.
XX
AC AAW14594;
XX
DT 06-FEB-1998 (first entry)
XX
DE LcnB bacteriocin with lactococcus protease PrtP signal peptide.
XX
KW Sec-dependent secretion system; bacteriocin; lactic acid bacteria;
KW fermentation; LcnB; P32 lactococcal promoter; lactococcal protease.
XX
OS Synthetic.
OS Lactococcus lactis.
XX
FH Key Location/Qualifiers
FT Peptide 1. .33
FT /label= Signal_peptide
FT Protein 34. .80
FT /label= LcnB
XX
PN WO9713863-A1.
XX
PD 17-APR-1997.
XX
PF 07-OCT-1996; 96WO-FR001560.
XX
PR 06-OCT-1995; 95FR-00011778.
XX
PA (SYST-) SYSTEMS BIO-IND.
XX
PI Fayard B, Kok J, Venema G, Bigret M, Prevots F;
XX
DR WPI; 1997-235899/21.
XX
PT Use of Sec-dependent system for secreting proteins usually secreted by
PT other methods - particularly for bacteriocin secretion in lactic acid
PT bacteria, for control of pathogenic bacteria and cell lysis.
XX
PS Example 2; Page 10; 21pp; French.
XX
CC The present sequence is derived from the LcnB bacteriocin and the
CC lactococcus protease PrtP signal peptide. The cDNA encoding this protein
CC was used in a new Sec-dependent secretion system for secreting proteins
CC normally secreted by a Sec-independent system. The DNA constructs for
CC this process comprise a promoter, signal sequence recognised by the Sec-
CC dependent system, sequence encoding mature and terminator. The method is
CC preferably used to secrete the bacteriocin LcnB of Lactococcus lactis,
CC especially in lactic acid bacteria. Bacteriocin- secreting cells are used
CC to prevent development of pathogenic bacteria and to lyse bacteria so
CC that the enzymes contained in them can take part in fermentations. Use of

```
CC   the Sec-dependent system provides high levels of secretion
XX
SQ   Sequence 80 AA;

  Query Match             91.7%;  Score 33;  DB 2;  Length 80;
  Best Local Similarity   50.0%;  Pred. No. 48;
  Matches    5;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 YXXXXGPXTW 11
              |    || ||
Db         37 YVMSAGPYTW 46


RESULT 6
ABB71807
ID   ABB71807 standard; protein; 2444 AA.
XX
AC   ABB71807;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster polypeptide SEQ ID NO 42213.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US009231.
XX
PR   23-MAR-2000; 2000US-0191637P.
PR   11-JUL-2000; 2000US-00614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
DR   N-PSDB; ABL15910.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signaling and cell-cell
PT   interactions.
XX
PS   Disclosure; SEQ ID NO 42213; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC   ABB72072). The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 2444 AA;

  Query Match             91.7%;  Score 33;  DB 4;  Length 2444;
  Best Local Similarity   50.0%;  Pred. No. 1.5e+03;
  Matches    5;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 YXXXXGPXTW 11
              |    || ||
Db       1646 YNTVSGPLTW 1655


RESULT 7
ADY54195
ID   ADY54195 standard; peptide; 17 AA.
XX
AC   ADY54195;
XX
DT   19-MAY-2005  (first entry)
XX
DE   Amino acid sequence of mutated EMP-1 #2.
XX
KW   cytostatic; anti-HIV; hypotensive; neuroprotective; cardiovascular-Gen.;
KW   nootropic; hepatotropic; virucide; antiinflammatory; immunosuppressive;
KW   antiallergic; antimicrobial; neuroleptic; gynecological; anorectic;
KW   antiarteriosclerotic; gastrointestinal-Gen.; endocrine-Gen; neoplasm;
KW   hematological disease; erythropoietin peptide mimetic; EPM;
KW   EPO mimetic peptide-1; EMP-1; multiple sclerosis; brain tumor; cancer;
KW   hepatitis; anemia; pregnancy; menstrual disorder; rheumatoid arthritis;
KW   AIDS; viral disease; metabolic disease; autoimmune disease;
KW   inflammatory disease; allergy; microbial infection;
KW   cardiovascular disease; genetic disease; neurodegenerative disease;
KW   hematopoietic cell disorder; endocrine disorder;
KW   gastrointestinal disease; hypertension; arterial sclerosis.
XX
OS   Synthetic.
XX
PN   WO2005021579-A2.
XX
PD   10-MAR-2005.
XX
PF   30-AUG-2004; 2004WO-US027949.
XX
PR   28-AUG-2003; 2003WO-US026818.
PR   10-MAR-2004; 2004US-0551552P.
XX
PA   (BIOR-) BIOREXIS PHARM CORP.
XX
PI   Sadeghi H,  Turner AJ;
XX
DR   WPI; 2005-214540/22.
XX
PT   Novel erythropoietin (EPO) peptide mimetic, having first modification of
PT   cysteine residue of EPO mimetic peptides (EMP)-1, to reduces disulfide
PT   bond formation, and second modification such that peptide exhibits EMP-1
PT   activity.
XX
PS   Example 2; SEQ ID NO 49; 158pp; English.
XX
CC   The specification describes an erythropoietin (EPO) peptide mimetic
CC   (EPM), comprising a modification of at least one cysteine residue of EPO
CC   mimetic peptide (EMP)-1 that substantially reduces disulfide bond
CC   formation, and a second modification such that the peptide exhibits EMP-1
CC   activity. The first modification comprises the deletion or substitution
CC   of at least one cysteine residue in EMP-1, and the second modification
CC   comprises the addition of a linker group that is covalently bonded to the
CC   C-terminal amino acid or N-terminal amino acid of EMP-1. EPM peptides of
CC   the invention are useful for treating or preventing diseases, such as
CC   multiple sclerosis, brain tumor, skin cancer, hepatitis B, hepatitis C,
CC   anemia, beta-thalassemia, pregnancy or menstrual disorders, rheumatoid
CC   arthritis, AIDS, cancer, viral disease, metabolic disease, obesity,
CC   autoimmune disease, inflammatory disease, allergy, graft-versus-host
CC   disease, systemic microbial infection, cardiovascular disease, psychosis,
CC   genetic diseases, neurodegenerative diseases, disorders of hematopoietic
CC   cells, diseases of the endocrine system or reproductive systems,
CC   gastrointestinal diseases, diabetes, asthma, or HIV infections,
CC   hypertension, hypercholesterolemia, arterial sclerosis, arthritis or
CC   Alzheimer's disease. The present sequence represents a mutated EMP-1,
CC   used to produce an EPM of the invention.
XX
SQ   Sequence 17 AA;

  Query Match             88.9%;  Score 32;  DB 9;  Length 17;
  Best Local Similarity   50.0%;  Pred. No. 16;
  Matches    5;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;
```

```
Qy          2 YXXXXGPXTW 11
              |    || ||
Db          4 YSGHFGPLTW 13
```

RESULT 8
ADY54194
ID ADY54194 standard; peptide; 17 AA.
XX
AC ADY54194;
XX
DT 19-MAY-2005 (first entry)
XX
DE Amino acid sequence of mutated EMP-1 #1.
XX
KW cytostatic; anti-HIV; hypotensive; neuroprotective; cardiovascular-Gen.;
KW nootropic; hepatotropic; virucide; antiinflammatory; immunosuppressive;
KW antiallergic; antimicrobial; neuroleptic; gynecological; anorectic;
KW antiarteriosclerotic; gastrointestinal-Gen.; endocrine-Gen; neoplasm;
KW hematological disease; erythropoietin peptide mimetic; EPM;
KW EPO mimetic peptide-1; EMP-1; multiple sclerosis; brain tumor; cancer;
KW hepatitis; anemia; pregnancy; menstrual disorder; rheumatoid arthritis;
KW AIDS; viral disease; metabolic disease; autoimmune disease;
KW inflammatory disease; allergy; microbial infection;
KW cardiovascular disease; genetic disease; neurodegenerative disease;
KW hematopoietic cell disorder; endocrine disorder;
KW gastrointestinal disease; hypertension; arterial sclerosis.
XX
OS Synthetic.
XX
PN WO2005021579-A2.
XX
PD 10-MAR-2005.
XX
PF 30-AUG-2004; 2004WO-US027949.
XX
PR 28-AUG-2003; 2003WO-US026818.
PR 10-MAR-2004; 2004US-0551552P.
XX
PA (BIOR-) BIOREXIS PHARM CORP.
XX
PI Sadeghi H, Turner AJ;
XX
DR WPI; 2005-214540/22.
XX
PT Novel erythropoietin (EPO) peptide mimetic, having first modification of
PT cysteine residue of EPO mimetic peptides (EMP)-1, to reduces disulfide
PT bond formation, and second modification such that peptide exhibits EMP-1
PT activity.
XX
PS Example 2; SEQ ID NO 48; 158pp; English.
XX
CC The specification describes an erythropoietin (EPO) peptide mimetic
CC (EPM), comprising a modification of at least one cysteine residue of EPO
CC mimetic peptide (EMP)-1 that substantially reduces disulfide bond
CC formation, and a second modification such that the peptide exhibits EMP-1
CC activity. The first modification comprises the deletion or substitution
CC of at least one cysteine residue in EMP-1, and the second modification
CC comprises the addition of a linker group that is covalently bonded to the
CC C-terminal amino acid or N-terminal amino acid of EMP-1. EPM peptides of
CC the invention are useful for treating or preventing diseases, such as
CC multiple sclerosis, brain tumor, skin cancer, hepatitis B, hepatitis C,
CC anemia, beta-thalassemia, pregnancy or menstrual disorders, rheumatoid
CC arthritis, AIDS, cancer, viral disease, metabolic disease, obesity,
CC autoimmune disease, inflammatory disease, allergy, graft-versus-host
CC disease, systemic microbial infection, cardiovascular disease, psychosis,
CC genetic diseases, neurodegenerative diseases, disorders of hematopoietic
CC cells, diseases of the endocrine system or reproductive systems,
CC gastrointestinal diseases, diabetes, asthma, or HIV infections,
CC hypertension, hypercholesterolemia, arterial sclerosis, arthritis or
CC Alzheimer's disease. The present sequence represents a mutated EMP-1,
CC used to produce an EPM of the invention.
XX
SQ Sequence 17 AA;

```
  Query Match             88.9%;  Score 32;  DB 9;  Length 17;
  Best Local Similarity   50.0%;  Pred. No. 16;
  Matches    5;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 YXXXXGPXTW 11
              |   . || ||
Db          4 YSGHFGPETW 13
```

RESULT 9
ADY54196
ID ADY54196 standard; peptide; 17 AA.
XX
AC ADY54196;
XX
DT 19-MAY-2005 (first entry)
XX
DE Amino acid sequence of mutated EMP-1 #3.
XX
KW cytostatic; anti-HIV; hypotensive; neuroprotective; cardiovascular-Gen.;
KW nootropic; hepatotropic; virucide; antiinflammatory; immunosuppressive;
KW antiallergic; antimicrobial; neuroleptic; gynecological; anorectic;
KW antiarteriosclerotic; gastrointestinal-Gen.; endocrine-Gen; neoplasm;
KW hematological disease; erythropoietin peptide mimetic; EPM;
KW EPO mimetic peptide-1; EMP-1; multiple sclerosis; brain tumor; cancer;
KW hepatitis; anemia; pregnancy; menstrual disorder; rheumatoid arthritis;
KW AIDS; viral disease; metabolic disease; autoimmune disease;
KW inflammatory disease; allergy; microbial infection;
KW cardiovascular disease; genetic disease; neurodegenerative disease;
KW hematopoietic cell disorder; endocrine disorder;
KW gastrointestinal disease; hypertension; arterial sclerosis.
XX
OS Synthetic.
XX
PN WO2005021579-A2.
XX
PD 10-MAR-2005.
XX
PF 30-AUG-2004; 2004WO-US027949.
XX
PR 28-AUG-2003; 2003WO-US026818.
PR 10-MAR-2004; 2004US-0551552P.
XX
PA (BIOR-) BIOREXIS PHARM CORP.
XX
PI Sadeghi H, Turner AJ;
XX
DR WPI; 2005-214540/22.
XX
PT Novel erythropoietin (EPO) peptide mimetic, having first modification of
PT cysteine residue of EPO mimetic peptides (EMP)-1, to reduces disulfide
PT bond formation, and second modification such that peptide exhibits EMP-1
PT activity.
XX
PS Example 2; SEQ ID NO 50; 158pp; English.
XX
CC The specification describes an erythropoietin (EPO) peptide mimetic
CC (EPM), comprising a modification of at least one cysteine residue of EPO
CC mimetic peptide (EMP)-1 that substantially reduces disulfide bond
CC formation, and a second modification such that the peptide exhibits EMP-1
CC activity. The first modification comprises the deletion or substitution
CC of at least one cysteine residue in EMP-1, and the second modification
CC comprises the addition of a linker group that is covalently bonded to the
CC C-terminal amino acid or N-terminal amino acid of EMP-1. EPM peptides of
CC the invention are useful for treating or preventing diseases, such as
CC multiple sclerosis, brain tumor, skin cancer, hepatitis B, hepatitis C,
CC anemia, beta-thalassemia, pregnancy or menstrual disorders, rheumatoid
CC arthritis, AIDS, cancer, viral disease, metabolic disease, obesity,
CC autoimmune disease, inflammatory disease, allergy, graft-versus-host
CC disease, systemic microbial infection, cardiovascular disease, psychosis,
CC genetic diseases, neurodegenerative diseases, disorders of hematopoietic

CC cells, diseases of the endocrine system or reproductive systems,
CC gastrointestinal diseases, diabetes, asthma, or HIV infections,
CC hypertension, hypercholesterolemia, arterial sclerosis, arthritis or
CC Alzheimer's disease. The present sequence represents a mutated EMP-1,
CC used to produce an EPM of the invention.
XX
SQ Sequence 17 AA;

Query Match 88.9%; Score 32; DB 9; Length 17;
Best Local Similarity 50.0%; Pred. No. 16;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 YXXXXGPXTW 11
| || ||
Db 4 YSGHFGPETW 13

RESULT 10
AAY26410
ID AAY26410 standard; peptide; 19 AA.
XX
AC AAY26410;
XX
DT 06-SEP-1999 (first entry)
XX
DE Erythropoietin receptor (EPO-R) binding peptide.
XX
KW Erythropoietin; EPO; receptor; EPO deficiency; renal failure; AIDS;
KW dialysis; anaemia; autoimmune disease; chronic inflammatory disease;
KW malignancy; beta-thalassaemia; cystic fibrosis; prematurity; blood loss;
KW spinal cord injury; aging; neoplastic disease; erythropoiesis; EPO-R.
XX
OS Synthetic.
XX
PN WO9640749-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US009810.
XX
PR 07-JUN-1995; 95US-00484631.
PR 07-JUN-1995; 95US-00484635.
XX
PA (JOHJ) JOHNSON & JOHNSON CORP.
PA (AFFY-) AFFYMAX TECHNOLOGIES NV.
XX
PI Wrighton NC, Dower WJ, Chang RS, Kashyap AK, Jolliffe LK;
PI Johnson D, Mulcahy L;
XX
DR WPI; 1997-052225/05.
XX
PT Erythropoietin receptor binding peptide - useful for treating disorders
PT characterised by deficiency of EPO, or low or defective red blood cell
PT population.
XX
PS Disclosure; Page 19; 95pp; English.
XX
CC The invention describes a peptide of 10-40 amino acid residues which
CC binds to erythropoietin (EPO) receptor and which includes the amino acid
CC sequence Cys-Xaa1-Xaa2-Gly-Pro-Xaa3-Thr-Trp-Xaa4-Cys, where Xaa1 = Arg,
CC His, Leu or Trp, Xaa2 = Met, Phe or Ile, Xaa3 = 1 of the 20 genetically
CC coded L-amino acids, and Xaa4 = Asp, Glu, Ile, Leu or Val. Optionally,
CC the peptide may be cyclised or dimerised. The peptide can be used to
CC treat a patient having a disorder characterised by a deficiency of EPO or
CC a low or defective red blood cell population. It can be used to treat end
CC stage renal failure or dialysis; anaemia associated with AIDS, autoimmune
CC disease, chronic inflammatory diseases or malignancy; beta-thalassaemia;
CC cystic fibrosis; early anaemia of prematurity; spinal cord injury; acute
CC blood loss; aging; and neoplastic disease states accompanied by abnormal
CC erythropoiesis. The peptides can also be used as reagents for detecting
CC EPO receptors on living cells, in biological fluids, in tissue
CC homogenates, etc. Sequences AAY26352-548 are representative peptides
CC falling within the above peptide motif and isolated during the affinity
CC selection process
XX
SQ Sequence 19 AA;

Query Match 88.9%; Score 32; DB 2; Length 19;
Best Local Similarity 50.0%; Pred. No. 18;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 YXXXXGPXTW 11
| || ||
Db 4 YMAHMGPITW 13

RESULT 11
AAY13631
ID AAY13631 standard; peptide; 19 AA.
XX
AC AAY13631;
XX
DT 06-SEP-1999 (first entry)
XX
DE Erythropoietin receptor (EPO-R) binding peptide.
XX
KW Erythropoietin; EPO; receptor; EPO deficiency; renal failure; AIDS;
KW dialysis; anaemia; autoimmune disease; chronic inflammatory disease;
KW malignancy; beta-thalassaemia; cystic fibrosis; prematurity; blood loss;
KW spinal cord injury; aging; neoplastic disease; erythropoiesis; EPO-R.
XX
OS Synthetic.
XX
PN WO9640749-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US009810.
XX
PR 07-JUN-1995; 95US-00484631.
PR 07-JUN-1995; 95US-00484635.
XX
PA (JOHJ) JOHNSON & JOHNSON CORP.
PA (AFFY-) AFFYMAX TECHNOLOGIES NV.
XX
PI Wrighton NC, Dower WJ, Chang RS, Kashyap AK, Jolliffe LK;
PI Johnson D, Mulcahy L;
XX
DR WPI; 1997-052225/05.
XX
PT Erythropoietin receptor binding peptide - useful for treating disorders
PT characterised by deficiency of EPO, or low or defective red blood cell
PT population.
XX
PS Claim 6; Page 68; 95pp; English.
XX
CC The invention describes a peptide of 10-40 amino acid residues which
CC binds to erythropoietin (EPO) receptor and which includes the amino acid
CC sequence Cys-Xaa1-Xaa2-Gly-Pro-Xaa3-Thr-Trp-Xaa4-Cys, where Xaa1 = Arg,
CC His, Leu or Trp, Xaa2 = Met, Phe or Ile, Xaa3 = 1 of the 20 genetically
CC coded L-amino acids, and Xaa4 = Asp, Glu, Ile, Leu or Val. Optionally,
CC the peptide may be cyclised or dimerised. The peptide can be used to
CC treat a patient having a disorder characterised by a deficiency of EPO or
CC a low or defective red blood cell population. It can be used to treat end
CC stage renal failure or dialysis; anaemia associated with AIDS, autoimmune
CC disease, chronic inflammatory diseases or malignancy; beta-thalassaemia;
CC cystic fibrosis; early anaemia of prematurity; spinal cord injury; acute
CC blood loss; aging; and neoplastic disease states accompanied by abnormal
CC erythropoiesis. The peptides can also be used as reagents for detecting
CC EPO receptors on living cells, in biological fluids, in tissue
CC homogenates, etc. Sequences AAY13624-661 represent specifc examples of
CC EPO-R binding peptides
XX
SQ Sequence 19 AA;

Query Match 88.9%; Score 32; DB 2; Length 19;

```
  Best Local Similarity   50.0%;  Pred. No. 18;
  Matches    5;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 YXXXXGPXTW 11
              |    || ||
Db          4 YMAHMGPITW 13


RESULT 12
AAW26967
ID   AAW26967 standard; peptide; 19 AA.
XX
AC   AAW26967;
XX
DT   10-NOV-1997  (first entry)
XX
DE   Monomer subunit of erythropoietin receptor binding dimer.
XX
KW   Monomer; erythropoietin; EPO; receptor; binding; dimer; activation;
KW   treatment; disorder; deficiency; low; defective; red blood cell;
KW   erythrocyte; population; cell surface; agonist; end stage; renal;
KW   failure; dialysis; anaemia; anemia; AIDS; chronic; inflammatory; disease;
KW   rheumatoid arthritis; bowel inflammation; autoimmune; transfusion.
XX
OS   Synthetic.
XX
PN   WO9640772-A2.
XX
PD   19-DEC-1996.
XX
PF   06-JUN-1996;   96WO-US009469.
XX
PR   07-JUN-1995;   95US-00484135.
XX
PA   (JOHJ ) JOHNSON & JOHNSON.
XX
PI   Johnson DL,  Zivin RA;
XX
DR   WPI; 1997-099920/09.
XX
PT   Activating cell surface receptors using peptide dimer agonists - also,
PT   new dimers of erythropoietin receptor binding peptide(s) useful for
PT   treating patient having disorder characterised by EPO deficiency.
XX
PS   Claim 6; Page 93; 110pp; English.
XX
CC   The present peptide is a monomer subunit of an erythropoietin (EPO)
CC   receptor binding dimer, which comprises 2 EPO receptor binding monomers
CC   of 10 to 40 amino acids, and activates or improves the bioactivity of the
CC   EPO cell surface receptor. The dimer can be used to treat disorders
CC   resulting from EPO deficiency by improving the activity of its cell
CC   surface receptor, e.g. end stage renal failure/dialysis, anaemia
CC   associated with AIDS or chronic inflammatory diseases such as rheumatoid
CC   arthritis and chronic bowel inflammation and autoimmune disease. It can
CC   also be used to boost the red cell count of a patient prior to surgery or
CC   as pretreatment to transfusion. The dimer peptide exhibits increased
CC   biological potency in vitro and in vivo relative to its component
CC   monomeric agonists. Dimerisation may also convert cell surface receptor
CC   antagonists into agonists
XX
SQ   Sequence 19 AA;

  Query Match             88.9%;  Score 32;  DB 2;  Length 19;
  Best Local Similarity   50.0%;  Pred. No. 18;
  Matches    5;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 YXXXXGPXTW 11
              |    || ||
Db          4 YMAHMGPITW 13


RESULT 13
AAB17931
ID   AAB17931 standard; peptide; 19 AA.
XX
AC   AAB17931;
XX
DT   31-OCT-2000  (first entry)
XX
DE   EPO-mimetic peptide sequence SEQ ID NO:1035.
XX
KW   Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW   autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW   immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW   inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW   cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW   vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW   thrombosis; pharmaceutical.
XX
OS   Synthetic.
XX
PN   WO200024782-A2.
XX
PD   04-MAY-2000.
XX
PF   25-OCT-1999;   99WO-US025044.
XX
PR   23-OCT-1998;   98US-0105371P.
PR   22-OCT-1999;   99US-00428082.
XX
PA   (AMGE-) AMGEN INC.
XX
PI   Feige U,  Liu C,  Cheetham J,  Boone TC;
XX
DR   WPI; 2000-350702/30.
XX
PT   Novel composition of matter comprising an Fc domain and pharmacologically
PT   active peptides, useful for treating cancer and autoimmune diseases.
XX
PS   Claim 13; Page 560; 608pp; English.
XX
CC   The present invention describes composition of matter (I) comprising an
CC   Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC   (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC   independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC   (L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC   P3, and P4 = are each independently sequences of pharmacologically active
CC   peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC   c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC   of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC   thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC   cells from the present invention can be used for producing pharmaceutical
CC   compositions. The compositions are useful for treating cancer, asthma,
CC   thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC   a Fab domain) can provide a longer half-life or incorporate functions
CC   such as Fc receptor binding, protein A binding, complement fixation, and
CC   possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC   AAB18003 represent nucleotide and amino acid sequences used in the
CC   exemplification of the present invention
XX
SQ   Sequence 19 AA;

  Query Match             88.9%;  Score 32;  DB 3;  Length 19;
  Best Local Similarity   50.0%;  Pred. No. 18;
  Matches    5;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 YXXXXGPXTW 11
              |    || ||
Db          4 YMAHMGPITW 13


RESULT 14
AAB17319
ID   AAB17319 standard; peptide; 19 AA.
XX
AC   AAB17319;
```

XX
DT 31-OCT-2000 (first entry)
XX
DE EPO-mimetic peptide sequence SEQ ID NO:423.
XX
KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist; MMP;
KW inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase; asthma;
KW thrombosis; pharmaceutical.
XX
OS Synthetic.
XX
PN WO200024782-A2.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-US025044.
XX
PR 23-OCT-1998; 98US-0105371P.
PR 22-OCT-1999; 99US-00428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and pharmacologically
PT active peptides, useful for treating cancer and autoimmune diseases.
XX
PS Claim 13; Page 345; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-
CC (L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2,
CC P3, and P4 = are each independently sequences of pharmacologically active
CC peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b,
CC c, d, e, and f = are each independently 0 or 1, provided that at least 1
CC of a and b is 1. The composition can have cytostatic, antiasthmatic,
CC thrombolytic and immunosuppressive activities. DNAs, vectors and host
CC cells from the present invention can be used for producing pharmaceutical
CC compositions. The compositions are useful for treating cancer, asthma,
CC thrombosis, or autoimmune diseases. The use of an Fc domain (rather than
CC a Fab domain) can provide a longer half-life or incorporate functions
CC such as Fc receptor binding, protein A binding, complement fixation, and
CC possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to
CC AAB18003 represent nucleotide and amino acid sequences used in the
CC exemplification of the present invention
XX
SQ Sequence 19 AA;

Query Match 88.9%; Score 32; DB 3; Length 19;
Best Local Similarity 50.0%; Pred. No. 18;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 YXXXXGPXTW 11
|| || ||
Db 4 YMAHMGPITW 13

RESULT 15
AAB13505
ID AAB13505 standard; peptide; 19 AA.
XX
AC AAB13505;
XX
DT 02-NOV-2000 (first entry)
XX
DE Erythropoietin derived peptide #6.
XX
KW Erythropoietin derivative; diagnostic agent; therapeutic agent.
XX
OS Unidentified.
XX
PN US6077939-A.
XX
PD 20-JUN-2000.
XX
PF 04-AUG-1997; 97US-00905310.
XX
PR 04-AUG-1997; 97US-00905310.
XX
PA (ORTH) ORTHO-MCNEIL PHARM INC.
XX
PI Wei Z, Ghosh-Dastidar P, Menon-Rudolph S;
XX
DR WPI; 2000-450981/39.
XX
PT Covalently binding a water-soluble polymer to the N-terminal alpha-carbon
PT atom of a polypeptide for producing therapeutic and diagnostic agents, by
PT contacting the transaminated polypeptide with the polymer to form a
PT hydrazone bond.
XX
PS Disclosure; Col 5; 21pp; English.
XX
CC The present sequence is a peptide derivative of erythropoietin. It can be
CC used in the method of the invention, which involves the production of
CC compositions comprising a polypeptide with a water soluble polymer
CC covalently bound to the N-terminal alpha carbon by a hydrazone or an
CC oxime (or the reduced version of each) bond. This is useful as previous
CC methods were more time consuming and labour-intensive, as well as being
CC less precise. The composition can be used to produce diagnostic and
CC therapeutic agents
XX
SQ Sequence 19 AA;

Query Match 88.9%; Score 32; DB 3; Length 19;
Best Local Similarity 50.0%; Pred. No. 18;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 YXXXXGPXTW 11
|| || ||
Db 4 YMAHMGPITW 13

Search completed: March 31, 2006, 17:39:31
Job time : 188 secs

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: March 31, 2006, 17:39:47 ; Search time 38 Seconds
(without alignments)
40.512 Million cell updates/sec

Title: US-10-609-217-419
Perfect score: 36
Sequence: 1 XYXXXXGPXTWXXXXX 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	33	91.7	68	2	B43940	lactococcin B prec
2	33	91.7	500	2	T49388	related to ascus d
3	31	86.1	310	2	T08078	carbonate dehydrat
4	31	86.1	334	2	T16772	hypothetical prote
5	31	86.1	395	2	B72478	hypothetical prote
6	31	86.1	476	2	T29025	hypothetical prote
7	31	86.1	1313	2	T29027	hypothetical prote
8	31	86.1	1506	2	T32909	hypothetical prote
9	30	83.3	668	2	T18635	hypothetical prote
10	29	80.6	227	2	C39925	hypothetical prote
11	29	80.6	237	2	G87286	conserved hypothet
12	29	80.6	246	2	B86784	hypothetical prote
13	29	80.6	263	2	T48742	hypothetical prote
14	29	80.6	279	2	AC3647	cellobiose phospho
15	29	80.6	332	2	B87356	sugar ABC transpor
16	29	80.6	334	2	JC6082	proximal sequence
17	29	80.6	371	2	T42623	probable sugar tra
18	29	80.6	433	2	T44587	cytochrome P450 ho
19	29	80.6	433	2	S63383	POP2 protein - yea
20	29	80.6	463	2	S36507	L2 protein - human
21	29	80.6	464	2	S36582	L2 protein - human
22	29	80.6	475	2	H84137	hypothetical prote
23	29	80.6	505	1	D70703	tldD homolog Rv231
24	29	80.6	506	2	D90207	conserved hypothet
25	29	80.6	511	2	A84537	probable sugar tra
26	29	80.6	511	2	H84536	probable sugar tra
27	29	80.6	536	2	D83622	arylsulfatase PA01
28	29	80.6	537	1	FOMVM7	gag polyprotein -
29	29	80.6	546	2	T40888	probable glucose t
30	29	80.6	565	2	T14732	probable beta-gluc
31	29	80.6	781	2	T49472	hormone-sensitive
32	29	80.6	821	2	B84509	probable Na/H anti
33	29	80.6	852	2	A34373	histidine-rich cal
34	29	80.6	1208	2	T00362	hypothetical prote
35	28	77.8	19	1	EWSMAN	ancovenin - Strept
36	28	77.8	60	2	S78724	protein YKL033w-a
37	28	77.8	62	2	B84394	hypothetical prote
38	28	77.8	68	2	S36976	hypothetical prote
39	28	77.8	93	2	T06470	probable chitinase
40	28	77.8	99	2	D75378	hypothetical prote
41	28	77.8	118	2	T17205	hypothetical prote
42	28	77.8	118	2	S59930	hypothetical prote
43	28	77.8	149	2	T26485	hypothetical prote
44	28	77.8	164	2	T04299	pathogenesis-relat
45	28	77.8	164	2	F83798	hypothetical prote

ALIGNMENTS

RESULT 1
B43940
lactococcin B precursor - Lactococcus lactis subsp. cremoris plasmid p9B4-6
C;Species: Lactococcus lactis subsp. cremoris
C;Date: 10-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: B43940
R;van Belkum, M.J.; Kok, J.; Venema, G.
Appl. Environ. Microbiol. 58, 572-577, 1992
A;Title: Cloning, sequencing, and expression in Escherichia coli of lcnB, a third bacter
A;Reference number: A43940; MUID:92304065; PMID:1610182
A;Accession: B43940
A;Molecule type: DNA
A;Residues: 1-68 <VAN>
A;Cross-references: UNIPROT:P35518; UNIPARC:UPI000012E2AA; GB:S38128; NID:g250436; PIDN:
A;Experimental source: subsp. cremoris, plasmid p9B4-6
A;Note: sequence extracted from NCBI backbone (NCBIN:106751, NCBIP:106754)
C;Keywords: antibiotic; bacteriocin
F;1-21/Domain: propeptide #status predicted <PRO>
F;22-68/Product: lactococcin B #status predicted <MAT>

Query Match 91.7%; Score 33; DB 2; Length 68;
Best Local Similarity 50.0%; Pred. No. 4.1;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy          2 YXXXXGPXTW 11
              |    || ||
Db         25 YVMSAGPYTW 34
```

RESULT 2
T49388
related to ascus development protein 3 [imported] - Neurospora crassa
N;Alternate names: protein B1D4.30
C;Species: Neurospora crassa
C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jul-2004
C;Accession: T49388
R;Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, May 2000
A;Reference number: Z25022
A;Accession: T49388
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-500 <SCH>
A;Cross-references: UNIPROT:Q9P679; UNIPARC:UPI000017B4DC; EMBL:AL355928; GSPDB:GN00116;
A;Experimental source: BAC clone B1D4; strain OR74A
C;Genetics:
A;Gene: NCSP:B1D4.30
A;Map position: 6
A;Introns: 16/2; 53/1; 130/3; 156/1; 394/3; 441/1

Query Match 91.7%; Score 33; DB 2; Length 500;
Best Local Similarity 50.0%; Pred. No. 29;

```
  Matches    5;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 YXXXXGPXTW 11
              |    || ||
Db        373 YSALLGPMTW 382


RESULT 3
T08078
carbonate dehydratase (EC 4.2.1.1) precursor, alpha type - Chlamydomonas reinhardtii
N;Alternate names: intracellular carbonic anhydrase
C;Species: Chlamydomonas reinhardtii
C;Date: 21-May-1999 #sequence_revision 21-May-1999 #text_change 09-Jul-2004
C;Accession: T08078
R;Karlsson, J.; Clarke, A.K.; Chen, Z.Y.; Hugghins, S.Y.; Park, Y.I.; Husic, H.D.; Moron
EMBO J. 17, 1208-1216, 1998
A;Title: A novel alpha-type carbonic anhydrase associated with the thylakoid membrane in
A;Reference number: Z16338; MUID:98151345; PMID:9482718
A;Accession: T08078
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-310 <KAR>
A;Cross-references: UNIPROT:Q39588; UNIPARC:UPI00000A34C2; EMBL:U40871; NID:g1655716; PI
C;Genetics:
A;Gene: CAH3
C;Function:
A;Description: catalyzes the reversible dissociation of carbonic acid to carbon dioxide
C;Superfamily: carbonate dehydratase; carbonic anhydrase homology
C;Keywords: carbon-oxygen lyase; hydro-lyase; thylakoid; zinc
F;1-72/Domain: transit peptide (chloroplast) #status predicted <TNP>
F;73-310/Product: carbonate dehydratase, alpha type #status predicted <MAT>
F;75-310/Domain: carbonic anhydrase homology <CAH>

  Query Match             86.1%;  Score 31;  DB 2;  Length 310;
  Best Local Similarity   50.0%;  Pred. No. 47;
  Matches    5;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 YXXXXGPXTW 11
              |    || ||
Db         77 YGEVAGPPTW 86


RESULT 4
T16772
hypothetical protein R173.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C;Accession: T16772
R;Geisel, C.
submitted to the EMBL Data Library, October 1995
A;Description: The sequence of C. elegans cosmid R173.
A;Reference number: Z18574
A;Accession: T16772
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-334 <GEI>
A;Cross-references: UNIPROT:Q10462; UNIPARC:UPI0000126DFF; EMBL:U39743; NID:g1049461; PI
C;Genetics:
A;Gene: CESP:R173.1
A;Introns: 4/3; 40/1; 72/3; 111/2; 164/3; 270/3
C;Superfamily: carbonate dehydratase; carbonic anhydrase homology

  Query Match             86.1%;  Score 31;  DB 2;  Length 334;
  Best Local Similarity   50.0%;  Pred. No. 50;
  Matches    5;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 YXXXXGPXTW 11
              |    || ||
Db         53 YDENNGPDTW 62


RESULT 5
B72478
hypothetical protein APE2466 - Aeropyrum pernix (strain K1)
C;Species: Aeropyrum pernix
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C;Accession: B72478
R;Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takaha
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Ku
DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyru
A;Reference number: A72450; MUID:99310339; PMID:10382966
A;Accession: B72478
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-395 <KAW>
A;Cross-references: UNIPROT:Q9Y918; UNIPARC:UPI000005E366; DDBJ:AP000064; NID:g5105945; I
A;Experimental source: strain K1
C;Genetics:
A;Gene: APE2466

  Query Match             86.1%;  Score 31;  DB 2;  Length 395;
  Best Local Similarity   50.0%;  Pred. No. 59;
  Matches    5;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 YXXXXGPXTW 11
              |    || ||
Db         56 YIDRMGPRTW 65


RESULT 6
T29025
hypothetical protein F53G12.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T29025
R;Wu, X.; Graves, T.
submitted to the EMBL Data Library, May 1997
A;Description: The sequence of C. elegans cosmid F53G12.
A;Reference number: Z20555
A;Accession: T29025
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-476 <WUX>
A;Cross-references: UNIPROT:Q9NH90; UNIPROT:O61213; UNIPARC:UPI000017BA07; EMBL:AF003139,
A;Experimental source: strain Bristol N2; clone F53G12
C;Genetics:
A;Gene: CESP:F53G12.2
A;Map position: 1
A;Introns: 53/3; 95/2; 165/3; 391/3; 435/3

  Query Match             86.1%;  Score 31;  DB 2;  Length 476;
  Best Local Similarity   50.0%;  Pred. No. 71;
  Matches    5;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 YXXXXGPXTW 11
              |    || ||
Db        257 YIKAVGPWTW 266


RESULT 7
T29027
hypothetical protein F53G12.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T29027
R;Wu, X.; Graves, T.
submitted to the EMBL Data Library, May 1997
A;Description: The sequence of C. elegans cosmid F53G12.
A;Reference number: Z20555
A;Accession: T29027
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1313 <WUX>
A;Cross-references: UNIPARC:UPI000017BA08; EMBL:AF003139; PIDN:AAB54159.1; GSPDB:GN00019
A;Experimental source: strain Bristol N2; clone F53G12
```

```
C;Genetics:
A;Gene: CESP:F53G12.3
A;Map position: 1
A;Introns: 123/1; 191/1; 275/1; 372/2; 443/3; 496/1; 538/1; 643/3; 743/1; 801/1; 842/2;

  Query Match             86.1%;  Score 31;  DB 2;  Length 1313;
  Best Local Similarity   50.0%;  Pred. No. 1.9e+02;
  Matches    5;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 YXXXXGPXTW 11
              |    || ||
Db       1167 YIKAVGPWTW 1176


RESULT 8
T32909
hypothetical protein F56C11.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C;Accession: T32909
R;Tin-Wollam, A.; Wohldmann, P.; Morris, M.
submitted to the EMBL Data Library, January 1998
A;Description: The sequence of C. elegans cosmid F56C11.
A;Reference number: Z21244
A;Accession: T32909
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1506 <TIN>
A;Cross-references: UNIPARC:UPI000017BA29; EMBL:AF043697; PIDN:AAB97555.1; GSPDB:GN00019
A;Experimental source: strain Bristol N2; clone F56C11
C;Genetics:
A;Gene: CESP:F56C11.1
A;Map position: 1
A;Introns: 21/1; 47/1; 100/1; 230/1; 298/1; 382/1; 479/2; 550/3; 599/1; 748/3; 848/1; 94

  Query Match             86.1%;  Score 31;  DB 2;  Length 1506;
  Best Local Similarity   50.0%;  Pred. No. 2.2e+02;
  Matches    5;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 YXXXXGPXTW 11
              |    || ||
Db       1274 YIKAVGPWTW 1283


RESULT 9
T18635
hypothetical protein B0019.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T18635
R;Kershaw, J.
submitted to the EMBL Data Library, November 1997
A;Reference number: Z19000
A;Accession: T18635
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-668 <WIL>
A;Cross-references: UNIPROT:Q9XXU5; UNIPARC:UPI0000163FC4; EMBL:AL008866; PIDN:CAA15509.
A;Experimental source: clone B0019
C;Genetics:
A;Gene: CESP:B0019.1
A;Map position: 1
A;Introns: 42/3; 103/1; 144/2; 195/3; 258/1; 304/1; 357/1; 404/2; 457/2; 474/1; 589/1; 6

  Query Match             83.3%;  Score 30;  DB 2;  Length 668;
  Best Local Similarity   40.0%;  Pred. No. 1.6e+02;
  Matches    4;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 YXXXXGPXTW 11
              |    ||  |
Db        335 YTSATGPLVW 344


RESULT 10
C39925
hypothetical protein 2 - equine arteritis virus
C;Species: equine arteritis virus
C;Date: 14-Feb-1992 #sequence_revision 14-Feb-1992 #text_change 09-Jul-2004
C;Accession: C39925
R;Den Boon, J.A.; Snijder, E.J.; Chirnside, E.D.; De Vries, A.A.F.; Horzinek, M.C.; Spaa
J. Virol. 65, 2910-2920, 1991
A;Title: Equine arteritis virus is not a togavirus but belongs to the coronaviruslike su
A;Reference number: A39925; MUID:91237805; PMID:1851863
A;Accession: C39925
A;Status: preliminary
A;Molecule type: genomic RNA
A;Residues: 1-227 <DEN>
A;Cross-references: UNIPROT:P28992; UNIPARC:UPI000011F47A; EMBL:X53459; NID:g62065; PIDN
C;Superfamily: equine arteritis virus hypothetical protein 2

  Query Match             80.6%;  Score 29;  DB 2;  Length 227;
  Best Local Similarity   80.0%;  Pred. No. 89;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          7 GPXTW 11
              || ||
Db        158 GPATW 162


RESULT 11
G87286
conserved hypothetical protein CC0304 [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C;Accession: G87286
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
 B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: G87286
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-237 <STO>
A;Cross-references: UNIPROT:Q9ABC6; UNIPARC:UPI00000C6FD0; GB:AE005673; NID:g13421447; P
C;Genetics:
A;Gene: CC0304

  Query Match             80.6%;  Score 29;  DB 2;  Length 237;
  Best Local Similarity   80.0%;  Pred. No. 93;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          7 GPXTW 11
              || ||
Db         64 GPATW 68


RESULT 12
B86784
hypothetical protein ynaE [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C;Species: Lactococcus lactis subsp. lactis
C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004
C;Accession: B86784
R;Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarme, K.; Weissenbach, J.; Ehrli
Genome Res. 11, 731-753, 2001
A;Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ss
A;Reference number: A86625; MUID:21235186; PMID:11337471
A;Accession: B86784
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-246 <STO>
A;Cross-references: UNIPROT:Q9CG36; UNIPARC:UPI00000C69F1; GB:AE005176; PID:g12724250; P
A;Experimental source: strain IL1403
C;Genetics:
A;Gene: ynaE
```

Query Match 80.6%; Score 29; DB 2; Length 246;
Best Local Similarity 80.0%; Pred. No. 97;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 7 GPXTW 11
|| ||
Db 115 GPSTW 119

RESULT 13
T48742
hypothetical protein 8D4.160 [imported] - Neurospora crassa
C;Species: Neurospora crassa
C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 19-May-2000
C;Accession: T48742
R;Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, April 2000
A;Reference number: Z24541
A;Accession: T48742
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-263 <SCH>
A;Cross-references: UNIPARC:UPI0000179476; EMBL:AL353819; GSPDB:GN00112; NCSP:8D4.160
A;Experimental source: cosmid contig 8D4; strain 74
C;Genetics:
A;Gene: NCSP:8D4.160
A;Map position: 2
A;Introns: 32/3; 76/1; 133/1
C;Superfamily: Neurospora crassa hypothetical protein 8D4.160

Query Match 80.6%; Score 29; DB 2; Length 263;
Best Local Similarity 80.0%; Pred. No. 1e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 7 GPXTW 11
|| ||
Db 252 GPSTW 256

RESULT 14
AC3647
cellobiose phosphotransferase system celC [imported] - Brucella melitensis (strain 16M)
C;Species: Brucella melitensis
C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C;Accession: AC3647
R;DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
.; Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitens
A;Reference number: AD3252; PMID:11756688
A;Accession: AC3647
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-279 <KUR>
A;Cross-references: UNIPROT:Q8YB01; UNIPARC:UPI0000058739; GB:AE008918; PIDN:AAL54342.1;
A;Experimental source: strain 16M
C;Genetics:
A;Gene: BMEII1100
A;Map position: II

Query Match 80.6%; Score 29; DB 2; Length 279;
Best Local Similarity 80.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 7 GPXTW 11
|| ||
Db 164 GPATW 168

RESULT 15
B87356
sugar ABC transporter, permease protein CC0861 [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C;Accession: B87356
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.E
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolona
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: B87356
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-332 <STO>
A;Cross-references: UNIPROT:Q9A9V0; UNIPARC:UPI00000C71C6; GB:AE005673; NID:g13422120; PI
C;Genetics:
A;Gene: CC0861
C;Superfamily: l-arabinose transport system permease araH

Query Match 80.6%; Score 29; DB 2; Length 332;
Best Local Similarity 80.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 7 GPXTW 11
|| ||
Db 104 GPATW 108

Search completed: March 31, 2006, 17:44:13
Job time : 38 secs

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: March 31, 2006, 17:36:37 ; Search time 233 Seconds
(without alignments)
48.448 Million cell updates/sec

Title: US-10-609-217-419
Perfect score: 36
Sequence: 1 XYXXXXGPXTWXXXXX 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	34	94.4	563	2	Q89XA5_BRAJA	Q89xa5 bradyrhizob
2	33	91.7	68	1	LCNB_LACLC	P35518 lactococcus
3	33	91.7	68	2	Q6TGC3_LACLA	Q6tgc3 lactococcus
4	33	91.7	220	2	Q4FPJ0_9RICK	Q4fpj0 candidatus
5	33	91.7	425	2	Q4T6Z9_TETNG	Q4t6z9 tetraodon n
6	33	91.7	460	2	Q89ZI0_BACTN	Q89zi0 bacteroides
7	33	91.7	493	2	Q9P679_NEUCR	Q9p679 neurospora
8	33	91.7	496	2	Q4HWT8_GIBZE	Q4hwt8 gibberella
9	33	91.7	500	2	Q7SC74_NEUCR	Q7sc74 neurospora
10	33	91.7	565	2	Q4IGZ6_GIBZE	Q4igz6 gibberella
11	33	91.7	940	2	Q95RI1_DROME	Q95ri1 drosophila
12	33	91.7	2618	2	Q7KTX8_DROME	Q7ktx8 drosophila
13	32	88.9	225	2	Q6IIU0_DROME	Q6iiu0 drosophila
14	32	88.9	285	2	Q81D50_BACCR	Q81d50 bacillus ce
15	32	88.9	332	2	Q9IK79_NPVST	Q9ik79 spodoptera
16	32	88.9	351	2	Q935B9_SALTI	Q935b9 salmonella
17	32	88.9	351	2	Q91BI7_NPVST	Q91bi7 spodoptera
18	32	88.9	474	2	Q6CV04_KLULA	Q6cv04 kluyveromyc
19	32	88.9	611	2	Q72HQ6_THET2	Q72hq6 thermus the
20	32	88.9	805	2	Q4Z454_PLABE	Q4z454 plasmodium
21	32	88.9	835	2	Q4XQ91_PLACH	Q4xq91 plasmodium
22	32	88.9	877	2	Q8IAY9_PLAF7	Q8iay9 plasmodium
23	32	88.9	877	2	Q7RGA4_PLAYO	Q7rga4 plasmodium
24	32	88.9	1123	2	Q7MXX2_PORGI	Q7mxx2 porphyromon
25	31	86.1	213	1	GIDB_RHIME	Q92kw3 rhizobium m
26	31	86.1	260	2	Q7ZYU6_XENLA	Q7zyu6 xenopus lae
27	31	86.1	260	2	Q8AVG8_XENLA	Q8avg8 xenopus lae
28	31	86.1	260	2	Q5FW20_XENTR	Q5fw20 xenopus tro
29	31	86.1	288	2	Q5DFH8_SCHJA	Q5dfh8 schistosoma
30	31	86.1	310	1	CAH5_CAEEL	Q10462 caenorhabdi
31	31	86.1	310	2	Q39588_CHLRE	Q39588 chlamydomon
32	31	86.1	395	2	Q9Y918_AERPE	Q9y918 aeropyrum p
33	31	86.1	445	2	Q8NCC2_HUMAN	Q8ncc2 homo sapien
34	31	86.1	465	2	Q66J34_XENLA	Q66j34 xenopus lae
35	31	86.1	482	2	Q4S9H8_TETNG	Q4s9h8 tetraodon n
36	31	86.1	497	2	Q8BTN2_MOUSE	Q8btn2 m mus muscu
37	31	86.1	507	1	GTR6_HUMAN	Q9ugq3 homo sapien
38	31	86.1	507	2	Q5SXD7_HUMAN	Q5sxd7 homo sapien
39	31	86.1	527	2	Q8MS78_DROME	Q8ms78 drosophila
40	31	86.1	579	2	Q9LD30_CRYCO	Q91d30 crypthecodi
41	31	86.1	719	2	Q7KWW5_DICDI	Q7kww5 dictyosteli
42	31	86.1	728	2	Q550Q3_DICDI	Q550q3 dictyosteli
43	31	86.1	752	2	Q8TMY4_METAC	Q8tmy4 methanosarc
44	31	86.1	1484	2	Q61UX8_CAEBR	Q61ux8 caenorhabdi
45	31	86.1	1497	2	Q9NH90_CAEEL	Q9nh90 caenorhabdi

ALIGNMENTS

RESULT 1
Q89XA5_BRAJA
ID Q89XA5_BRAJA PRELIMINARY; PRT; 563 AA.
AC Q89XA5;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Bll0409 protein.
GN OrderedLocusNames=bll0409;
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=USDA 110;
RX MEDLINE=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA Sasamoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110.";
RL DNA Res. 9:189-197(2002).
DR EMBL; BA000040; BAC45674.1; -; Genomic_DNA.
DR GO; GO:0015036; F:disulfide oxidoreductase activity; IEA.
DR GO; GO:0004497; F:monooxygenase activity; IEA.
DR GO; GO:0006725; P:aromatic compound metabolism; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001327; FAD_pyr_redox.
DR InterPro; IPR002938; Moxy_FAD_binding.
DR InterPro; IPR000205; NAD_BS.
DR InterPro; IPR001100; Pyr_redox.
DR InterPro; IPR003042; Rng_mnoxygenase.
DR Pfam; PF01494; FAD_binding_3; 1.
DR PRINTS; PR00368; FADPNR.
DR PRINTS; PR00411; PNDRDTASEI.
DR PRINTS; PR00420; RNGMNOXGNASE.
KW Complete proteome.
SQ SEQUENCE 563 AA; 61540 MW; 8A002A8636CF7268 CRC64;

Query Match 94.4%; Score 34; DB 2; Length 563;
Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy          2 YXXXXGPXTW 11
              |    || ||
Db        117 YSATGGPDTW 126
```

RESULT 2
LCNB_LACLC
ID LCNB_LACLC STANDARD; PRT; 68 AA.

```
AC   P35518;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Bacteriocin lactococcin B precursor (LCN-B).
GN   Name=lcnB;
OS   Lactococcus lactis subsp. cremoris (Streptococcus cremoris).
OG   Plasmid p9B4-6.
OC   Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX   NCBI_TaxID=1359;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=9B4;
RX   MEDLINE=92304065; PubMed=1610182;
RA   van Belkum M.J., Kok J., Venema G.;
RT   "Cloning, sequencing, and expression in Escherichia coli of lcnB, a
RT   third bacteriocin determinant from the lactococcal bacteriocin plasmid
RT   p9B4-6.";
RL   Appl. Environ. Microbiol. 58:572-577(1992).
RN   [2]
RP   MUTAGENESIS OF CYS-45.
RX   MEDLINE=97039852; PubMed=8885398;
RA   Venema K., Dost M.H., Venema G., Kok J.;
RT   "Mutational analysis and chemical modification of Cys24 of lactococcin
RT   B, a bacteriocin produced by Lactococcus lactis.";
RL   Microbiology 142:2825-2830(1996).
CC   -!- FUNCTION: Kills Lactococci by dissipating the membrane potential
CC       of the cells.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; S38128; AAB22372.1; -; Genomic_DNA.
DR   PIR; B43940; B43940.
DR   InterPro; IPR010133; Bacteriocin_sig.
DR   InterPro; IPR007464; Lactococcin.
DR   Pfam; PF04369; Lactococcin; 1.
DR   TIGRFAMs; TIGR01847; bacteriocin_sig; 1.
KW   Antibiotic; Antimicrobial; Bacteriocin; Plasmid; Transmembrane.
FT   PROPEP        1     21
FT   CHAIN        22     68       Bacteriocin lactococcin B.
FT   MUTAGEN      45     45       C->A,D,E,F,G,I,L,M,N,P,Q,S,T,V,Y,W: No
FT                                loss of activity.
FT   MUTAGEN      45     45       C->R,K: Loss of activity.
SQ   SEQUENCE   68 AA;  7632 MW;  18382310AC880678 CRC64;

  Query Match             91.7%;  Score 33;  DB 1;  Length 68;
  Best Local Similarity   50.0%;  Pred. No. 31;
  Matches    5;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 YXXXXGPXTW 11
              |    || ||
Db         25 YVMSAGPYTW 34


RESULT 3
Q6TGC3_LACLA
ID   Q6TGC3_LACLA PRELIMINARY;      PRT;    68 AA.
AC   Q6TGC3;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   LcnB.
GN   Name=lcnB;
OS   Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OG   Plasmid BGMN1-5.
OC   Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX   NCBI_TaxID=1360;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Gajic O., Buist G., Topisirovic L., Venema G., Kok J., Kuipers O.P.;
RL   Submitted (SEP-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY422080; AAQ97214.1; -; Genomic_DNA.
DR   GO; GO:0005576; C:extracellular region; IEA.
DR   GO; GO:0042742; P:defense response to bacteria; IEA.
DR   InterPro; IPR010133; Bacteriocin_sig.
DR   InterPro; IPR007464; Lactococcin.
DR   Pfam; PF04369; Lactococcin; 1.
DR   TIGRFAMs; TIGR01847; bacteriocin_sig; 1.
KW   Plasmid.
SQ   SEQUENCE   68 AA;  7632 MW;  18382310AC880678 CRC64;

  Query Match             91.7%;  Score 33;  DB 2;  Length 68;
  Best Local Similarity   50.0%;  Pred. No. 31;
  Matches    5;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 YXXXXGPXTW 11
              |    || ||
Db         25 YVMSAGPYTW 34


RESULT 4
Q4FPJ0_9RICK
ID   Q4FPJ0_9RICK PRELIMINARY;      PRT;   220 AA.
AC   Q4FPJ0;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Hypothetical protein.
GN   ORFNames=SAR11_0075;
OS   Candidatus Pelagibacter ubique HTCC1062.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC   SAR11 cluster; Candidatus Pelagibacter.
OX   NCBI_TaxID=335992;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=HTCC1062;
RA   Giovannoni S.J., Tripp H.J., Givan S.A., Podar M., Vergin K.L.,
RA   Baptista D., Bibbs L., Eads J., Richardson T.H., Noordweir M.,
RA   Rappe M.S., Short J., Carrington J.C., Mathur E.J.;
RT   "Genome Streamlining in a Cosmopolitan Oceanic Bacterium.";
RL   Submitted (JUL-2005) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; CP000084; AAZ20899.1; -; Genomic_DNA.
KW   Hypothetical protein.
SQ   SEQUENCE   220 AA;  25476 MW;  A5D24C4475A1F796 CRC64;

  Query Match             91.7%;  Score 33;  DB 2;  Length 220;
  Best Local Similarity   50.0%;  Pred. No. 94;
  Matches    5;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 YXXXXGPXTW 11
              |    || ||
Db        182 YMFTSGPNTW 191


RESULT 5
Q4T6Z9_TETNG
ID   Q4T6Z9_TETNG PRELIMINARY;      PRT;   425 AA.
AC   Q4T6Z9;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Chromosome undetermined SCAF8419, whole genome shotgun sequence.
DE   (Fragment).
GN   ORFNames=GSTENG00006010001;
OS   Tetraodon nigroviridis (Green puffer).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC   Tetradontoidea; Tetraodontidae; Tetraodon.
OX   NCBI_TaxID=99883;
RN   [1]
```

```
RP   NUCLEOTIDE SEQUENCE.
RA   Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA   Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA   Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA   Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA   Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA   Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA   Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA   Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA   Kellis M., Volff JN., Guigo R., Zody M.C., Mesirov J.,
RA   Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA   Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA   Wincker P., Lander E.S., Weissenbach J., Roest Crollius H.;
RT   "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT   the early vertebrate proto-karyotype.";
RL   Nature 431:946-957(2004).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RG   Genoscope; Whitehead Institute Centre for Genome Research;
RL   Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
DR   EMBL; CAAE01008419; CAF91333.1; -; Genomic_DNA.
DR   InterPro; IPR007114; MFS.
DR   InterPro; IPR005828; Sub_transporter.
DR   InterPro; IPR003663; Sugar_transpt.
DR   Pfam; PF00083; Sugar_tr; 1.
DR   PRINTS; PR00171; SUGRTRNSPORT.
DR   PROSITE; PS50850; MFS; 1.
KW   Transmembrane; Transport.
FT   NON_TER       1      1
FT   NON_TER     425    425
SQ   SEQUENCE   425 AA;  46001 MW;  504081DC04569CAC CRC64;

  Query Match             91.7%;  Score 33;  DB 2;  Length 425;
  Best Local Similarity   50.0%;  Pred. No. 1.7e+02;
  Matches    5;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 YXXXXGPXTW 11
              |    || ||
Db        327 YSAGFGPMTW 336


RESULT 6
Q89ZI0_BACTN
ID   Q89ZI0_BACTN PRELIMINARY;      PRT;   460 AA.
AC   Q89ZI0;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Xylose/H+ symporter.
GN   OrderedLocusNames=BT4397;
OS   Bacteroides thetaiotaomicron.
OC   Bacteria; Bacteroidetes; Bacteroidetes (class); Bacteroidales;
OC   Bacteroidaceae; Bacteroides.
OX   NCBI_TaxID=818;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=VPI-5482 / ATCC 29148;
RX   MEDLINE=22550858; PubMed=12663928; DOI=10.1126/science.1080029;
RA   Xu J., Bjursell M.K., Himrod J., Deng S., Carmichael L.K.,
RA   Chiang H.C., Hooper L.V., Gordon J.I.;
RT   "A genomic view of the human-Bacteroides thetaiotaomicron symbiosis.";
RL   Science 299:2074-2076(2003).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
DR   EMBL; AE016945; AAO79502.1; -; Genomic_DNA.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0005351; F:sugar porter activity; IEA.
DR   GO; GO:0005215; F:transporter activity; IEA.
DR   GO; GO:0008643; P:carbohydrate transport; IEA.
DR   InterPro; IPR007114; MFS.
DR   InterPro; IPR005828; Sub_transporter.
DR   InterPro; IPR005829; Sug_transporter.
DR   InterPro; IPR003663; Sugar_transpt.
DR   Pfam; PF00083; Sugar_tr; 1.
DR   PRINTS; PR00171; SUGRTRNSPORT.
DR   TIGRFAMs; TIGR00879; SP; 1.
DR   PROSITE; PS50850; MFS; 1.
DR   PROSITE; PS00216; SUGAR_TRANSPORT_1; 2.
DR   PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW   Complete proteome; Transmembrane; Transport.
SQ   SEQUENCE   460 AA;  50207 MW;  5FCE2A30680A5C4C CRC64;

  Query Match             91.7%;  Score 33;  DB 2;  Length 460;
  Best Local Similarity   50.0%;  Pred. No. 1.9e+02;
  Matches    5;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 YXXXXGPXTW 11
              |    || ||
Db        369 YAMSLGPITW 378


RESULT 7
Q9P679_NEUCR
ID   Q9P679_NEUCR PRELIMINARY;      PRT;   493 AA.
AC   Q9P679;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Related to ascus development protein 3.
GN   Name=B1D4.030;
OS   Neurospora crassa.
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC   Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX   NCBI_TaxID=5141;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Schulte U., Aign V., Hoheisel J., Brandt P., Fartmann B., Holland R.,
RA   Nyakatura G., Mewes H.W., Mannhaupt G.;
RL   Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RA   German Neurospora genome project;
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
DR   EMBL; AL355928; CAB91291.2; -; Genomic_DNA.
DR   PIR; T49388; T49388.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0005351; F:sugar porter activity; IEA.
DR   GO; GO:0005215; F:transporter activity; IEA.
DR   GO; GO:0008643; P:carbohydrate transport; IEA.
DR   InterPro; IPR007114; MFS.
DR   InterPro; IPR005828; Sub_transporter.
DR   InterPro; IPR005829; Sug_transporter.
DR   InterPro; IPR003663; Sugar_transpt.
DR   Pfam; PF00083; Sugar_tr; 1.
DR   PRINTS; PR00171; SUGRTRNSPORT.
DR   PROSITE; PS50850; MFS; 1.
DR   PROSITE; PS00216; SUGAR_TRANSPORT_1; 2.
KW   Transmembrane; Transport.
SQ   SEQUENCE   493 AA;  54562 MW;  89010B4A27A5A87B CRC64;

  Query Match             91.7%;  Score 33;  DB 2;  Length 493;
  Best Local Similarity   50.0%;  Pred. No. 2e+02;
  Matches    5;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 YXXXXGPXTW 11
              |    || ||
Db        366 YSALLGPMTW 375


RESULT 8
```

```
Q4HWT8_GIBZE
ID   Q4HWT8_GIBZE PRELIMINARY;      PRT;   496 AA.
AC   Q4HWT8;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Hypothetical protein.
GN   ORFNames=FG10570.1;
OS   Gibberella zeae PH-1.
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC   Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
OX   NCBI_TaxID=229533;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=PH-1;
RA   Birren B., Nusbaum C., Abouelleil A., Allen N., Anderson S.,
RA   Arachchi H.M., Barna N., Bastien V., Bloom T., Boguslavkiy L.,
RA   Boukhgalter B., Butler J., Calvo S.E., Camarata J., Chang J.,
RA   Choepel Y., Collymore A., Cook A., Cooke P., Corum B., DeArellano K.,
RA   Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
RA   Erickson J., Faro S., Ferreira P., FitzGerald M., Gage D., Galagan J.,
RA   Gardyna S., Gnerre S., Graham L., Grand-Pierre N., Hafez N.,
RA   Hagopian D., Hagos B., Hall J., Horton L., Hulme W., Iliev I.,
RA   Jaffe D., Johnson R., Jones C., Kamal M., Kamat A., Karatas A.,
RA   Kells C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
RA   Ma L.-J., Mabbitt R., MacLean C., Macdonald P., Major J., Manning J.,
RA   Matthews C., Mauceli E., McCarthy M., Meldrim J., Meneus L.,
RA   Mihova T., Mlenga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
RA   Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neil D.,
RA   Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,
RA   Rachupka A., Ramasamy U., Raymond C., Retta R., Rise C., Rogov P.,
RA   Roman J., Schauer S., Schupback R., Seaman S., Severy P., Smirnov S.,
RA   Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,
RA   Talamas J., Tesfaye S., Theodore J., Topham K., Travers M.,
RA   Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,
RA   Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
RA   Lander E.;
RT   "Fusarium graminearum genome sequence.";
RL   Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; AACM01000442; EAA68456.1; -; Genomic_DNA.
KW   Hypothetical protein.
SQ   SEQUENCE   496 AA;  54315 MW;  551D9ED9AE843E57 CRC64;

  Query Match             91.7%;  Score 33;  DB 2;  Length 496;
  Best Local Similarity   50.0%;  Pred. No. 2e+02;
  Matches    5;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 YXXXXGPXTW 11
              |    || ||
Db        376 YSALLGPMTW 385


RESULT 9
Q7SC74_NEUCR
ID   Q7SC74_NEUCR PRELIMINARY;      PRT;   500 AA.
AC   Q7SC74;
DT   01-MAR-2004 (TrEMBLrel. 26, Created)
DT   01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Related to ascus development protein 3.
GN   Name=NCU05350.1;
OS   Neurospora crassa.
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC   Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX   NCBI_TaxID=5141;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=OR74A;
RA   Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
RA   Jaffe D., FitzHugh W., Ma L.-J., Smirnov S., Purcell S., Rehman B.,
RA   Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
RA   Qui D., Ianakiev P., Pedersen D., Nelson M., Washburne M.,
RA   Selitrennikoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,
RA   Kothe G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,
RA   Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S.,
RA   Kamal M., Kamvysselis M., Mauceli E., Bielke C., Rudd S., Frishman D.,
RA   Krystofova S., Rasmussen C., Metzenberg R.L., Perkins D.D., Kroken S.,
RA   Cogoni C., Macino G., Catcheside D., Li W., Pratt R.J., Osmani S.A.,
RA   DeSouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,
RA   Yarden O., Plamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,
RA   Natvig D.O., Alex L.A., Mannhaupt G., Ebbole D.J., Freitag M.,
RA   Paulsen I., Sachs M.S., Lander E.S., Nusbaum C., Birren B.;
RT   "The Genome Sequence of the Filamentous Fungus Neurospora crassa.";
RL   Nature 0:0-0(2003).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; AABX01000137; EAA34105.1; -; Genomic_DNA.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0005351; F:sugar porter activity; IEA.
DR   GO; GO:0005215; F:transporter activity; IEA.
DR   GO; GO:0008643; P:carbohydrate transport; IEA.
DR   InterPro; IPR007114; MFS.
DR   InterPro; IPR005828; Sub_transporter.
DR   InterPro; IPR005829; Sug_transporter.
DR   InterPro; IPR003663; Sugar_transpt.
DR   Pfam; PF00083; Sugar_tr; 1.
DR   PRINTS; PR00171; SUGRTRNSPORT.
DR   PROSITE; PS50850; MFS; 1.
DR   PROSITE; PS00216; SUGAR_TRANSPORT_1; 2.
KW   Transmembrane; Transport.
SQ   SEQUENCE   500 AA;  55248 MW;  E37DFA56E99FF0A0 CRC64;

  Query Match             91.7%;  Score 33;  DB 2;  Length 500;
  Best Local Similarity   50.0%;  Pred. No. 2e+02;
  Matches    5;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 YXXXXGPXTW 11
              |    || ||
Db        373 YSALLGPMTW 382


RESULT 10
Q4IGZ6_GIBZE
ID   Q4IGZ6_GIBZE PRELIMINARY;      PRT;   565 AA.
AC   Q4IGZ6;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Hypothetical protein.
GN   ORFNames=FG03512.1;
OS   Gibberella zeae PH-1.
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC   Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
OX   NCBI_TaxID=229533;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=PH-1;
RA   Birren B., Nusbaum C., Abouelleil A., Allen N., Anderson S.,
RA   Arachchi H.M., Barna N., Bastien V., Bloom T., Boguslavkiy L.,
RA   Boukhgalter B., Butler J., Calvo S.E., Camarata J., Chang J.,
RA   Choepel Y., Collymore A., Cook A., Cooke P., Corum B., DeArellano K.,
RA   Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
RA   Erickson J., Faro S., Ferreira P., FitzGerald M., Gage D., Galagan J.,
RA   Gardyna S., Gnerre S., Graham L., Grand-Pierre N., Hafez N.,
RA   Hagopian D., Hagos B., Hall J., Horton L., Hulme W., Iliev I.,
RA   Jaffe D., Johnson R., Jones C., Kamal M., Kamat A., Karatas A.,
RA   Kells C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
RA   Ma L.-J., Mabbitt R., MacLean C., Macdonald P., Major J., Manning J.,
RA   Matthews C., Mauceli E., McCarthy M., Meldrim J., Meneus L.,
RA   Mihova T., Mlenga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
```

```
RA   Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neil D.,
RA   Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,
RA   Rachupka A., Ramasamy U., Raymond C., Retta R., Rise C., Rogov P.,
RA   Roman J., Schauer S., Schupback R., Seaman S., Severy P., Smirnov S.,
RA   Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,
RA   Talamas J., Tesfaye S., Theodore J., Topham K., Travers M.,
RA   Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,
RA   Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
RA   Lander E.;
RT   "Fusarium graminearum genome sequence.";
RL   Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; AACM01000159; EAA72478.1; -; Genomic_DNA.
KW   Hypothetical protein.
SQ   SEQUENCE   565 AA;  61090 MW;  FD54AA196600C62A CRC64;

  Query Match             91.7%;  Score 33;  DB 2;  Length 565;
  Best Local Similarity   50.0%;  Pred. No. 2.3e+02;
  Matches    5;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 YXXXXGPXTW 11
              |    || ||
Db        395 YASSWGPVTW 404


RESULT 11
Q95RI1_DROME
ID   Q95RI1_DROME PRELIMINARY;      PRT;   940 AA.
AC   Q95RI1;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   LD28662p.
GN   Name=skd; Synonyms=pap;
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Berkeley;
RA   Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA   Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA   Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA   Nunoo J., Pacleb J., Paragas V., Park S., Phouanenavong S., Wan K.,
RA   Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY061361; AAL28909.1; -; mRNA.
DR   FlyBase; FBgn0003415; skd.
DR   GO; GO:0005634; C:nucleus; IDA.
DR   GO; GO:0045165; P:cell fate commitment; IGI.
DR   GO; GO:0009790; P:embryonic development; IMP.
DR   GO; GO:0045498; P:sex comb development; IGI.
DR   InterPro; IPR009401; TRAP_240kDa.
DR   Pfam; PF06333; TRAP_240kDa; 1.
SQ   SEQUENCE   940 AA;  104682 MW;  62593A53251BB38D CRC64;

  Query Match             91.7%;  Score 33;  DB 2;  Length 940;
  Best Local Similarity   50.0%;  Pred. No. 3.7e+02;
  Matches    5;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 YXXXXGPXTW 11
              |    || ||
Db        142 YNTVSGPLTW 151


RESULT 12
Q7KTX8_DROME
ID   Q7KTX8_DROME PRELIMINARY;      PRT;  2618 AA.
AC   Q7KTX8;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   CG9936-PC, isoform C (Cg9936-pd, isoform d) (Pap/DTRAP240) (Jlkjlkj)
DE   (Transcriptional coactivator blind spot).
GN   Name=skd; Synonyms=bli, pap; ORFNames=CG9936;
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22426065; PubMed=12537568;
RA   Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA   Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA   George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,
RA   Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA   Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA   Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT   "Finishing a whole-genome shotgun: release 3 of the Drosophila
RT   melanogaster euchromatic genome sequence.";
RL   Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22426070; PubMed=12537573;
RA   Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J.W., Svirskas R.,
RA   Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA   Ashburner M., Celniker S.E.;
RT   "The transposable elements of the Drosophila melanogaster euchromatin:
RT   a genomics perspective.";
RL   Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).
```

```
RN   [4]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22426069; PubMed=12537572;
RA   Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA   Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA   Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA   Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA   Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA   Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA   Lewis S.E.;
RT   "Annotation of the Drosophila melanogaster euchromatic genome: a
RT   systematic review.";
RL   Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN   [5]
RP   NUCLEOTIDE SEQUENCE.
RG   Berkeley Drosophila Genome Project;
RA   Celniker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,
RA   Hoskins R., Stapleton M., Pacleb J., Park S., Svirskas R., Smith E.,
RA   Yu C., Rubin G.;
RT   "Drosophila melanogaster release 4 sequence.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   NUCLEOTIDE SEQUENCE.
RG   FlyBase;
RL   Submitted (MAR-2005) to the EMBL/GenBank/DDBJ databases.
RN   [7]
RP   NUCLEOTIDE SEQUENCE.
RA   Boube M., Faucher C., Joulia L., Cribbs D.L., Bourbon H.M.;
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
RN   [8]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=21098949; PubMed=11171343;
RA   Treisman J.;
RT   "Drosophila homologues of the transcriptional coactivation complex
RT   subunits TRAP240 and TRAP230 are required for identical processes in
RT   eye-antennal disc development.";
RL   Development 128:603-615(2001).
RN   [9]
RP   NUCLEOTIDE SEQUENCE.
RA   Treisman J.E.;
RL   Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
RN   [10]
RP   NUCLEOTIDE SEQUENCE.
RA   Nairz K., Hafen E.;
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE003593; AAN12148.1; -; Genomic_DNA.
DR   EMBL; AF227214; AAF43021.1; -; mRNA.
DR   EMBL; AF227215; AAF43172.1; -; Genomic_DNA.
DR   EMBL; AF324425; AAG48327.1; -; mRNA.
DR   EMBL; AF226855; AAF36691.1; -; mRNA.
DR   Ensembl; CG9936; Drosophila melanogaster.
DR   InterPro; IPR009401; TRAP_240kDa.
DR   Pfam; PF06333; TRAP_240kDa; 1.
SQ   SEQUENCE   2618 AA;  280021 MW;  735A8A502076844E CRC64;

  Query Match             91.7%;  Score 33;  DB 2;  Length 2618;
  Best Local Similarity   50.0%;  Pred. No. 9.5e+02;
  Matches    5;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 YXXXXGPXTW 11
              |    || ||
Db       1820 YNTVSGPLTW 1829


RESULT 13
Q6IIU0_DROME
ID   Q6IIU0_DROME PRELIMINARY;      PRT;   225 AA.
AC   Q6IIU0;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   HDC17109.
GN   ORFNames=HDC17109;
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   PubMed=14709175; DOI=10.1186/gb-2003-5-1-r3;
RA   Hild M., Beckmann B., Haas S.A., Koch B., Solovyev V., Busold C.,
RA   Fellenberg K., Boutros M., Vingron M., Sauer F., Hoheisel J.D.,
RA   Paro R.;
RT   "An integrated gene annotation and transcriptional profiling approach
RT   towards the full gene content of the Drosophila genome.";
RL   Genome Biol. 5:RESEARCH0003.1-RESEARCH0003.17(2003).
CC   -!- MISCELLANEOUS: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ third party annotation (TPA) entry.
DR   EMBL; BK002976; DAA03176.1; -; Genomic_DNA.
SQ   SEQUENCE   225 AA;  25168 MW;  FC643BAAD34AFA59 CRC64;

  Query Match             88.9%;  Score 32;  DB 2;  Length 225;
  Best Local Similarity   50.0%;  Pred. No. 1.5e+02;
  Matches    5;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 YXXXXGPXTW 11
              |    || ||
Db         60 YRLTLGPLTW 69


RESULT 14
Q81D50_BACCR
ID   Q81D50_BACCR PRELIMINARY;      PRT;   285 AA.
AC   Q81D50;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   DTDP-glucose 4,6-dehydratase (EC 4.2.1.46).
GN   OrderedLocusNames=BC2530;
OS   Bacillus cereus (strain ATCC 14579 / DSM 31).
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;
OC   Bacillus cereus group.
OX   NCBI_TaxID=226900;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22608415; PubMed=12721630; DOI=10.1038/nature01582;
RA   Ivanova N., Sorokin A., Anderson I., Galleron N., Candelon B.,
RA   Kapatral V., Bhattacharyya A., Reznik G., Mikhailova N., Lapidus A.,
RA   Chu L., Mazur M., Goltsman E., Larsen N., D'Souza M., Walunas T.,
RA   Grechkin Y., Pusch G., Haselkorn R., Fonstein M., Ehrlich S.D.,
RA   Overbeek R., Kyrpides N.C.;
RT   "Genome sequence of Bacillus cereus and comparative analysis with
RT   Bacillus anthracis.";
RL   Nature 423:87-91(2003).
DR   EMBL; AE017006; AAP09490.1; -; Genomic_DNA.
DR   GO; GO:0008460; F:dTDP-glucose 4,6-dehydratase activity; IEA.
DR   GO; GO:0016829; F:lyase activity; IEA.
DR   InterPro; IPR001509; Epimerase_Dh.
DR   Pfam; PF01370; Epimerase; 1.
KW   Complete proteome; Lyase.
SQ   SEQUENCE   285 AA;  31263 MW;  4C744F7C20128452 CRC64;

  Query Match             88.9%;  Score 32;  DB 2;  Length 285;
  Best Local Similarity   50.0%;  Pred. No. 1.9e+02;
  Matches    5;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 YXXXXGPXTW 11
              |    || ||
Db        163 YGTLYGPGTW 172


RESULT 15
Q9IK79_NPVST
ID   Q9IK79_NPVST PRELIMINARY;      PRT;   332 AA.
AC   Q9IK79;
```

```
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   GP37 protein.
OS   Spodoptera litura multicapsid nucleopolyhedrovirus (SpltMNPV).
OC   Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC   Nucleopolyhedrovirus.
OX   NCBI_TaxID=46242;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Li C., Pang Y., Yan Q.;
RL   Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF216301; AAF72586.1; -; Genomic_DNA.
DR   HSSP; Q862M4; 1AAR.
DR   SMR; Q9IK79; 1-76.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   GO; GO:0006464; P:protein modification; IEA.
DR   InterPro; IPR004302; Chitin_binding_3.
DR   InterPro; IPR000626; Ubiquitin.
DR   Pfam; PF03067; Chitin_bind_3; 1.
DR   Pfam; PF00240; ubiquitin; 1.
DR   PRINTS; PR00348; UBIQUITIN.
DR   SMART; SM00213; UBQ; 1.
DR   PROSITE; PS00299; UBIQUITIN_1; 1.
DR   PROSITE; PS50053; UBIQUITIN_2; 1.
SQ   SEQUENCE   332 AA;  37558 MW;  02345EFA8E52AD12 CRC64;

  Query Match             88.9%;  Score 32;  DB 2;  Length 332;
  Best Local Similarity   50.0%;  Pred. No. 2.2e+02;
  Matches    5;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 YXXXXGPXTW 11
              |    || ||
Db        245 YDADGGPLTW 254


Search completed: March 31, 2006, 17:43:31
Job time : 236 secs
```

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: March 31, 2006, 16:09:06 ; Search time 61.6915 Seconds
(without alignments)
113.955 Million cell updates/sec

Title: US-10-609-217-420
Perfect score: 54
Sequence: 1 XYXCXXGPXTWXCXXX 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	52	96.3	21	9	ADU91978	Adu91978 EPO-R ago
2	51	94.4	20	2	AAY13696	Aay13696 Erythropo
3	51	94.4	20	2	AAY13650	Aay13650 Erythropo
4	51	94.4	20	2	AAY13728	Aay13728 Erythropo
5	51	94.4	20	2	AAY13687	Aay13687 Erythropo
6	51	94.4	20	2	AAW27001	Aaw27001 Monomer s
7	51	94.4	20	2	AAW27010	Aaw27010 Monomer s
8	51	94.4	20	5	AAU74480	Aau74480 Human ery
9	51	94.4	22	2	AAY13709	Aay13709 Erythropo
10	51	94.4	22	2	AAY26491	Aay26491 Erythropo
11	51	94.4	22	2	AAY26355	Aay26355 Erythropo
12	51	94.4	22	2	AAW27023	Aaw27023 Monomer s
13	51	94.4	22	9	ADU91911	Adu91911 EPO-R ago
14	51	94.4	23	9	ADU91986	Adu91986 EPO-R ago
15	51	94.4	23	9	ADU91945	Adu91945 EPO-R ago
16	51	94.4	24	2	AAY26424	Aay26424 Erythropo
17	51	94.4	24	2	AAY26431	Aay26431 Erythropo
18	51	94.4	133	7	ADJ73535	Adj73535 Erythropo
19	50	92.6	12	2	AAY13658	Aay13658 Erythropo
20	50	92.6	12	5	AAU74473	Aau74473 Human ery
21	50	92.6	13	2	AAY13636	Aay13636 Erythropo
22	50	92.6	13	2	AAW26971	Aaw26971 Monomer s
23	50	92.6	13	3	AAB17324	Aab17324 EPO-mimet
24	50	92.6	13	3	AAB17935	Aab17935 EPO-mimet
25	50	92.6	13	3	AAB13509	Aab13509 Erythropo
26	50	92.6	13	5	ABB72848	Abb72848 Erythropo
27	50	92.6	13	5	ABB72833	Abb72833 Erythropo
28	50	92.6	13	5	AAU74464	Aau74464 Human ery
29	50	92.6	13	7	ADJ72572	Adj72572 EPO mimet
30	50	92.6	13	7	ADJ72587	Adj72587 EPO mimet
31	50	92.6	13	8	ADK15878	Adk15878 EMP20 pep
32	50	92.6	13	8	ADJ52223	Adj52223 CH1 delet
33	50	92.6	13	8	ADJ52208	Adj52208 CH1 delet
34	50	92.6	13	8	ADJ51185	Adj51185 CH1 delet
35	50	92.6	13	8	ADJ51170	Adj51170 CH1 delet
36	50	92.6	13	8	ADL24422	Adl24422 EPO activ
37	50	92.6	13	9	ADZ44417	Adz44417 Erythropo
38	50	92.6	13	9	ADZ44428	Adz44428 Erythropo
39	50	92.6	14	2	AAY13654	Aay13654 Erythropo
40	50	92.6	14	5	AAU74472	Aau74472 Human ery
41	50	92.6	16	2	AAY13653	Aay13653 Erythropo
42	50	92.6	16	5	AAU74471	Aau74471 Human ery
43	50	92.6	17	9	ADU91963	Adu91963 EPO-R ago
44	50	92.6	17	9	ADU91877	Adu91877 EPO-R ago
45	50	92.6	17	9	ADU91957	Adu91957 EPO-R ago

ALIGNMENTS

RESULT 1
ADU91978
ID ADU91978 standard; peptide; 21 AA.
XX
AC ADU91978;
XX
DT 10-FEB-2005 (first entry)
XX
DE EPO-R agonist SEQ ID NO 119.
XX
KW erythropoietin receptor; EPO-R; erythropoietin; renal failure;
KW autoimmune disease; cystic fibrosis; anemia; inflammation;
KW spinal cord injury; aging; neurological disease; nephrotropic;
KW antianemic; immunosuppressive; CNS-Gen.; neuroprotective;
KW respiratory-Gen.; antiinflammatory; vulnerary; nootropic; cytostatic;
KW hemostatic; cyclic.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "Acetylated residue"
FT Disulfide-bond 7. .16
FT Modified-site 21
FT /note= "C-terminal amide"
XX
PN WO2004101611-A2.
XX
PD 25-NOV-2004.
XX
PF 12-MAY-2004; 2004WO-US014886.
XX
PR 12-MAY-2003; 2003US-0470245P.
XX
PA (AFFY-) AFFYMAX INC.
XX
PI Yin K, Holmes C, Lalonde G, Balu P, Schatz PJ, Tumelty D;
XX
DR WPI; 2005-039329/04.
XX
PT New peptide comprising specified sequence of amino acid is erythropoietin
PT receptor agonist useful for treating e.g. anemia, beta-thalassemia, renal
PT disorders.
XX
PS Disclosure; SEQ ID NO 119; 83pp; English.
XX
CC This invention describes a novel peptide which is an erythropoietin

CC receptor (EPO-R) activator. The peptide forms a dimer comprising a
CC linking moiety connecting two peptide chains composed of ADU91861. The N-
CC terminal of the peptide is acetylated. The EPO-R activator further
CC comprises at least one water soluble polymer, preferably polyethylene
CC glycol (PEG) covalently bound to the peptide and a spacer moiety. The
CC products of the invention are used for treating disorders associated with
CC deficiency of erythropoietin or low or defective red blood cell
CC population, end stage renal failure or dialysis, anemia associated with
CC AIDS, autoimmune disease or malignancy, beta-thalassemia, cystic
CC fibrosis, early anemia of prematurity, anemia associated with chronic
CC inflammatory disease, spinal cord injury, acute blood loss, aging and
CC neoplastic disease states accompanied by abnormal erythropoiesis. The
CC peptide compounds are potent agonists of erythropoietin receptor and have
CC nephrotropic, antianemic, immunosuppressive, CNS-Gen., neuroprotective,
CC respiratory-Gen., antiinflammatory, vulnerary, nootropic, cytostatic and
CC hemostatic activity. This sequence represents a peptide which acts as an
CC erythropoietin receptor (EPO-R) agonist.
XX
SQ Sequence 21 AA;

Query Match 96.3%; Score 52; DB 9; Length 21;
Best Local Similarity 58.3%; Pred. No. 0.1;
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 YXCXXGPXTWXC 13
| | || || |
Db 5 YSCQMGPTTWTC 16

RESULT 2
AAY13696
ID AAY13696 standard; peptide; 20 AA.
XX
AC AAY13696;
XX
DT 06-SEP-1999 (first entry)
XX
DE Erythropoietin receptor (EPO-R) binding peptide.
XX
KW Erythropoietin; EPO; receptor; EPO deficiency; renal failure; AIDS;
KW dialysis; anaemia; autoimmune disease; chronic inflammatory disease;
KW malignancy; beta-thalassaemia; cystic fibrosis; prematurity; blood loss;
KW spinal cord injury; aging; neoplastic disease; erythropoiesis; EPO-R.
XX
OS Synthetic.
XX
PN WO9640749-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US009810.
XX
PR 07-JUN-1995; 95US-00484631.
PR 07-JUN-1995; 95US-00484635.
XX
PA (JOHJ) JOHNSON & JOHNSON CORP.
PA (AFFY-) AFFYMAX TECHNOLOGIES NV.
XX
PI Wrighton NC, Dower WJ, Chang RS, Kashyap AK, Jolliffe LK;
PI Johnson D, Mulcahy L;
XX
DR WPI; 1997-052225/05.
XX
PT Erythropoietin receptor binding peptide - useful for treating disorders
PT characterised by deficiency of EPO, or low or defective red blood cell
PT population.
XX
PS Disclosure; Fig 2; 95pp; English.
XX
CC The invention describes a peptide of 10-40 amino acid residues which
CC binds to erythropoietin (EPO) receptor and which includes the amino acid
CC sequence Cys-Xaa1-Xaa2-Gly-Pro-Xaa3-Thr-Trp-Xaa4-Cys, where Xaa1 = Arg,
CC His, Leu or Trp, Xaa2 = Met, Phe or Ile, Xaa3 = 1 of the 20 genetically
CC coded L-amino acids, and Xaa4 = Asp, Glu, Ile, Leu or Val. Optionally,
CC the peptide may be cyclised or dimerised. The peptide can be used to
CC treat a patient having a disorder characterised by a deficiency of EPO or
CC a low or defective red blood cell population. It can be used to treat end
CC stage renal failure or dialysis; anaemia associated with AIDS, autoimmune
CC disease, chronic inflammatory diseases or malignancy; beta-thalassaemia;
CC cystic fibrosis; early anaemia of prematurity; spinal cord injury; acute
CC blood loss; aging; and neoplastic disease states accompanied by abnormal
CC erythropoiesis. The peptides can also be used as reagents for detecting
CC EPO receptors on living cells, in biological fluids, in tissue
CC homogenates, etc. Sequences AAY13662-735 are representative peptides of
CC the invention
XX
SQ Sequence 20 AA;

Query Match 94.4%; Score 51; DB 2; Length 20;
Best Local Similarity 58.3%; Pred. No. 0.14;
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 YXCXXGPXTWXC 13
| | || || |
Db 4 YSCRMGPTTWEC 15

RESULT 3
AAY13650
ID AAY13650 standard; peptide; 20 AA.
XX
AC AAY13650;
XX
DT 06-SEP-1999 (first entry)
XX
DE Erythropoietin receptor (EPO-R) binding peptide.
XX
KW Erythropoietin; EPO; receptor; EPO deficiency; renal failure; AIDS;
KW dialysis; anaemia; autoimmune disease; chronic inflammatory disease;
KW malignancy; beta-thalassaemia; cystic fibrosis; prematurity; blood loss;
KW spinal cord injury; aging; neoplastic disease; erythropoiesis; EPO-R.
XX
OS Synthetic.
XX
PN WO9640749-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US009810.
XX
PR 07-JUN-1995; 95US-00484631.
PR 07-JUN-1995; 95US-00484635.
XX
PA (JOHJ) JOHNSON & JOHNSON CORP.
PA (AFFY-) AFFYMAX TECHNOLOGIES NV.
XX
PI Wrighton NC, Dower WJ, Chang RS, Kashyap AK, Jolliffe LK;
PI Johnson D, Mulcahy L;
XX
DR WPI; 1997-052225/05.
XX
PT Erythropoietin receptor binding peptide - useful for treating disorders
PT characterised by deficiency of EPO, or low or defective red blood cell
PT population.
XX
PS Claim 6; Page 68; 95pp; English.
XX
CC The invention describes a peptide of 10-40 amino acid residues which
CC binds to erythropoietin (EPO) receptor and which includes the amino acid
CC sequence Cys-Xaa1-Xaa2-Gly-Pro-Xaa3-Thr-Trp-Xaa4-Cys, where Xaa1 = Arg,
CC His, Leu or Trp, Xaa2 = Met, Phe or Ile, Xaa3 = 1 of the 20 genetically
CC coded L-amino acids, and Xaa4 = Asp, Glu, Ile, Leu or Val. Optionally,
CC the peptide may be cyclised or dimerised. The peptide can be used to
CC treat a patient having a disorder characterised by a deficiency of EPO or
CC a low or defective red blood cell population. It can be used to treat end
CC stage renal failure or dialysis; anaemia associated with AIDS, autoimmune

CC disease, chronic inflammatory diseases or malignancy; beta-thalassaemia;
CC cystic fibrosis; early anaemia of prematurity; spinal cord injury; acute
CC blood loss; aging; and neoplastic disease states accompanied by abnormal
CC erythropoiesis. The peptides can also be used as reagents for detecting
CC EPO receptors on living cells, in biological fluids, in tissue
CC homogenates, etc. Sequences AAY13624-661 represent specifc examples of
CC EPO-R binding peptides
XX
SQ Sequence 20 AA;

Query Match 94.4%; Score 51; DB 2; Length 20;
Best Local Similarity 58.3%; Pred. No. 0.14;
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 YXCXXGPXTWXC 13
Db 4 YSCHFGPATWVC 15

RESULT 4
AAY13728
ID AAY13728 standard; peptide; 20 AA.
XX
AC AAY13728;
XX
DT 06-SEP-1999 (first entry)
XX
DE Erythropoietin receptor (EPO-R) binding peptide.
XX
KW Erythropoietin; EPO; receptor; EPO deficiency; renal failure; AIDS;
KW dialysis; anaemia; autoimmune disease; chronic inflammatory disease;
KW malignancy; beta-thalassaemia; cystic fibrosis; prematurity; blood loss;
KW spinal cord injury; aging; neoplastic disease; erythropoiesis; EPO-R.
XX
OS Synthetic.
XX
PN WO9640749-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US009810.
XX
PR 07-JUN-1995; 95US-00484631.
PR 07-JUN-1995; 95US-00484635.
XX
PA (JOHJ) JOHNSON & JOHNSON CORP.
PA (AFFY-) AFFYMAX TECHNOLOGIES NV.
XX
PI Wrighton NC, Dower WJ, Chang RS, Kashyap AK, Jolliffe LK;
PI Johnson D, Mulcahy L;
XX
DR WPI; 1997-052225/05.
XX
PT Erythropoietin receptor binding peptide - useful for treating disorders
PT characterised by deficiency of EPO, or low or defective red blood cell
PT population.
XX
PS Disclosure; Fig 2; 95pp; English.
XX
CC The invention describes a peptide of 10-40 amino acid residues which
CC binds to erythropoietin (EPO) receptor and which includes the amino acid
CC sequence Cys-Xaa1-Xaa2-Gly-Pro-Xaa3-Thr-Trp-Xaa4-Cys, where Xaa1 = Arg,
CC His, Leu or Trp, Xaa2 = Met, Phe or Ile, Xaa3 = 1 of the 20 genetically
CC coded L-amino acids, and Xaa4 = Asp, Glu, Ile, Leu or Val. Optionally,
CC the peptide may be cyclised or dimerised. The peptide can be used to
CC treat a patient having a disorder characterised by a deficiency of EPO or
CC a low or defective red blood cell population. It can be used to treat end
CC stage renal failure or dialysis; anaemia associated with AIDS, autoimmune
CC disease, chronic inflammatory diseases or malignancy; beta-thalassaemia;
CC cystic fibrosis; early anaemia of prematurity; spinal cord injury; acute
CC blood loss; aging; and neoplastic disease states accompanied by abnormal
CC erythropoiesis. The peptides can also be used as reagents for detecting
CC EPO receptors on living cells, in biological fluids, in tissue
CC homogenates, etc. Sequences AAY13662-735 are representative peptides of
CC the invention
XX
SQ Sequence 20 AA;

Query Match 94.4%; Score 51; DB 2; Length 20;
Best Local Similarity 58.3%; Pred. No. 0.14;
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 YXCXXGPXTWXC 13
Db 4 YACRMGPTTWVC 15

RESULT 5
AAY13687
ID AAY13687 standard; peptide; 20 AA.
XX
AC AAY13687;
XX
DT 06-SEP-1999 (first entry)
XX
DE Erythropoietin receptor (EPO-R) binding peptide.
XX
KW Erythropoietin; EPO; receptor; EPO deficiency; renal failure; AIDS;
KW dialysis; anaemia; autoimmune disease; chronic inflammatory disease;
KW malignancy; beta-thalassaemia; cystic fibrosis; prematurity; blood loss;
KW spinal cord injury; aging; neoplastic disease; erythropoiesis; EPO-R.
XX
OS Synthetic.
XX
PN WO9640749-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US009810.
XX
PR 07-JUN-1995; 95US-00484631.
PR 07-JUN-1995; 95US-00484635.
XX
PA (JOHJ) JOHNSON & JOHNSON CORP.
PA (AFFY-) AFFYMAX TECHNOLOGIES NV.
XX
PI Wrighton NC, Dower WJ, Chang RS, Kashyap AK, Jolliffe LK;
PI Johnson D, Mulcahy L;
XX
DR WPI; 1997-052225/05.
XX
PT Erythropoietin receptor binding peptide - useful for treating disorders
PT characterised by deficiency of EPO, or low or defective red blood cell
PT population.
XX
PS Disclosure; Fig 2; 95pp; English.
XX
CC The invention describes a peptide of 10-40 amino acid residues which
CC binds to erythropoietin (EPO) receptor and which includes the amino acid
CC sequence Cys-Xaa1-Xaa2-Gly-Pro-Xaa3-Thr-Trp-Xaa4-Cys, where Xaa1 = Arg,
CC His, Leu or Trp, Xaa2 = Met, Phe or Ile, Xaa3 = 1 of the 20 genetically
CC coded L-amino acids, and Xaa4 = Asp, Glu, Ile, Leu or Val. Optionally,
CC the peptide may be cyclised or dimerised. The peptide can be used to
CC treat a patient having a disorder characterised by a deficiency of EPO or
CC a low or defective red blood cell population. It can be used to treat end
CC stage renal failure or dialysis; anaemia associated with AIDS, autoimmune
CC disease, chronic inflammatory diseases or malignancy; beta-thalassaemia;
CC cystic fibrosis; early anaemia of prematurity; spinal cord injury; acute
CC blood loss; aging; and neoplastic disease states accompanied by abnormal
CC erythropoiesis. The peptides can also be used as reagents for detecting
CC EPO receptors on living cells, in biological fluids, in tissue
CC homogenates, etc. Sequences AAY13662-735 are representative peptides of
CC the invention
XX
SQ Sequence 20 AA;

```
  Query Match             94.4%;  Score 51;  DB 2;  Length 20;
  Best Local Similarity   58.3%;  Pred. No. 0.14;
  Matches    7;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 YXCXXGPXTWXC 13
              | |  || || |
Db          4 YSCRMGPTTWVC 15


RESULT 6
AAW27001
ID   AAW27001 standard; peptide; 20 AA.
XX
AC   AAW27001;
XX
DT   11-NOV-1997  (first entry)
XX
DE   Monomer subunit of erythropoietin receptor binding dimer.
XX
KW   Monomer; erythropoietin; EPO; receptor; binding; dimer; activation;
KW   treatment; disorder; deficiency; low; defective; red blood cell;
KW   erythrocyte; population; cell surface; agonist; end stage; renal;
KW   failure; dialysis; anaemia; anemia; AIDS; chronic; inflammatory; disease;
KW   rheumatoid arthritis; bowel inflammation; autoimmune; transfusion.
XX
OS   Synthetic.
XX
PN   WO9640772-A2.
XX
PD   19-DEC-1996.
XX
PF   06-JUN-1996;   96WO-US009469.
XX
PR   07-JUN-1995;   95US-00484135.
XX
PA   (JOHJ ) JOHNSON & JOHNSON.
XX
PI   Johnson DL,  Zivin RA;
XX
DR   WPI; 1997-099920/09.
XX
PT   Activating cell surface receptors using peptide dimer agonists - also,
PT   new dimers of erythropoietin receptor binding peptide(s) useful for
PT   treating patient having disorder characterised by EPO deficiency.
XX
PS   Disclosure; Fig 9; 110pp; English.
XX
CC   The present peptide is a specific example of a claimed generic monomer
CC   subunit of an erythropoietin (EPO) receptor binding dimer, which
CC   comprises 2 EPO receptor binding monomers of 10 to 40 amino acids, and
CC   activates or improves the bioactivity of the EPO cell surface receptor.
CC   The dimer can be used to treat disorders resulting from EPO deficiency by
CC   improving the activity of its cell surface receptor, e.g. end stage renal
CC   failure/dialysis, anaemia associated with AIDS or chronic inflammatory
CC   diseases such as rheumatoid arthritis and chronic bowel inflammation and
CC   autoimmune disease. It can also be used to boost the red cell count of a
CC   patient prior to surgery or as pretreatment to transfusion. The dimer
CC   peptide exhibits increased biological potency in vitro and in vivo
CC   relative to its component monomeric agonists. Dimerisation may also
CC   convert cell surface receptor antagonists into agonists
XX
SQ   Sequence 20 AA;

  Query Match             94.4%;  Score 51;  DB 2;  Length 20;
  Best Local Similarity   58.3%;  Pred. No. 0.14;
  Matches    7;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 YXCXXGPXTWXC 13
              | |  || || |
Db          4 YSCRMGPTTWVC 15


RESULT 7
AAW27010
ID   AAW27010 standard; peptide; 20 AA.
XX
AC   AAW27010;
XX
DT   11-NOV-1997  (first entry)
XX
DE   Monomer subunit of erythropoietin receptor binding dimer.
XX
KW   Monomer; erythropoietin; EPO; receptor; binding; dimer; activation;
KW   treatment; disorder; deficiency; low; defective; red blood cell;
KW   erythrocyte; population; cell surface; agonist; end stage; renal;
KW   failure; dialysis; anaemia; anemia; AIDS; chronic; inflammatory; disease;
KW   rheumatoid arthritis; bowel inflammation; autoimmune; transfusion.
XX
OS   Synthetic.
XX
PN   WO9640772-A2.
XX
PD   19-DEC-1996.
XX
PF   06-JUN-1996;   96WO-US009469.
XX
PR   07-JUN-1995;   95US-00484135.
XX
PA   (JOHJ ) JOHNSON & JOHNSON.
XX
PI   Johnson DL,  Zivin RA;
XX
DR   WPI; 1997-099920/09.
XX
PT   Activating cell surface receptors using peptide dimer agonists - also,
PT   new dimers of erythropoietin receptor binding peptide(s) useful for
PT   treating patient having disorder characterised by EPO deficiency.
XX
PS   Disclosure; Fig 9; 110pp; English.
XX
CC   The present peptide is a specific example of a claimed generic monomer
CC   subunit of an erythropoietin (EPO) receptor binding dimer, which
CC   comprises 2 EPO receptor binding monomers of 10 to 40 amino acids, and
CC   activates or improves the bioactivity of the EPO cell surface receptor.
CC   The dimer can be used to treat disorders resulting from EPO deficiency by
CC   improving the activity of its cell surface receptor, e.g. end stage renal
CC   failure/dialysis, anaemia associated with AIDS or chronic inflammatory
CC   diseases such as rheumatoid arthritis and chronic bowel inflammation and
CC   autoimmune disease. It can also be used to boost the red cell count of a
CC   patient prior to surgery or as pretreatment to transfusion. The dimer
CC   peptide exhibits increased biological potency in vitro and in vivo
CC   relative to its component monomeric agonists. Dimerisation may also
CC   convert cell surface receptor antagonists into agonists
XX
SQ   Sequence 20 AA;

  Query Match             94.4%;  Score 51;  DB 2;  Length 20;
  Best Local Similarity   58.3%;  Pred. No. 0.14;
  Matches    7;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 YXCXXGPXTWXC 13
              | |  || || |
Db          4 YSCRMGPTTWEC 15


RESULT 8
AAU74480
ID   AAU74480 standard; peptide; 20 AA.
XX
AC   AAU74480;
XX
DT   09-APR-2002  (first entry)
XX
DE   Human erythropoietin neuroprotective peptide EMP-11.
XX
KW   Neuroprotective peptide; human; erythropoietin; neurotoxicity; anaemia;
```

KW erythropoietin receptor; EPO receptor; neurodegeneration; prion disease;
KW neurological damage; neurodegenerative disorder; neurological disorder;
KW psychiatric disorder; blood loss; renal failure; cancer; HIV;
KW human immunodeficiency virus; haematology; autoimmune disease;
KW inflammatory disorder; infectious disease; EMP-11.
XX
OS Homo sapiens.
XX
PN WO200191780-A1.
XX
PD 06-DEC-2001.
XX
PF 23-MAY-2001; 2001WO-US016654.
XX
PR 26-MAY-2000; 2000US-0207654P.
XX
PA (ORTH) ORTHO-MCNEIL PHARM INC.
XX
PI Smith-Swintosky V, Renzi M, Plata-Salaman C, Jolliffe L;
PI Farrell F, Johnson DL;
XX
DR WPI; 2002-114307/15.
XX
PT Treating patients having condition mediated by neurotoxicity,
PT neurodegeneration or neurological damage, involves administering to
PT patient a peptide comprising monomeric peptides that bind to
PT erythropoietin receptor.
XX
PS Claim 14; Page 41; 75pp; English.
XX
CC The invention relates to a method for treating a patient with a condition
CC mediated by neurotoxicity, neurodegeneration or neurological damage,
CC involving administering a peptide comprising one or more monomeric
CC peptides that bind to the human erythropoietin (EPO) receptor. The method
CC is useful for treating acute and chronic neurodegenerative disorders
CC including cerebral ischaemia or infarction, Alzheimer's disease, Pick's
CC disease, diffuse Lewy body disease, Shy-Drager syndrome, amyotrophic
CC lateral sclerosis, Huntington's disease, Parkinson's disease, Gilles De
CC La Tourette's disease, Tay-Sach's disease, and prion diseases including
CC Creutzfeldt-Jakob and Kuru, neurological and psychiatric manifestations
CC associated with peripheral diseases including blood loss of any kind,
CC renal failure, conditions associated with anaemia, and neurological and
CC neuropsychiatric manifestations including haematological and non-
CC haematological malignancies/cancer, symptoms or complications in patients
CC receiving chemotherapy, inflammatory and infectious disorders such as
CC human immunodeficiency viral infections, and chronic systemic autoimmune
CC diseases such as systemic lupus erythematosus. The method is also useful
CC for prevention of plexopathies and neuropathies. This sequence represents
CC a human erythropoietin neuroprotective peptide of the invention
XX
SQ Sequence 20 AA;

```
  Query Match             94.4%;  Score 51;  DB 5;  Length 20;
  Best Local Similarity   58.3%;  Pred. No. 0.14;
  Matches    7;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 YXCXXGPXTWXC 13
              | |  || || |
Db          4 YSCHFGPATWVC 15
```

RESULT 9
AAY13709
ID AAY13709 standard; peptide; 22 AA.
XX
AC AAY13709;
XX
DT 06-SEP-1999 (first entry)
XX
DE Erythropoietin receptor (EPO-R) binding peptide.
XX
KW Erythropoietin; EPO; receptor; EPO deficiency; renal failure; AIDS;
KW dialysis; anaemia; autoimmune disease; chronic inflammatory disease;
KW malignancy; beta-thalassaemia; cystic fibrosis; prematurity; blood loss;
KW spinal cord injury; aging; neoplastic disease; erythropoiesis; EPO-R.
XX
OS Synthetic.
XX
PN WO9640749-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US009810.
XX
PR 07-JUN-1995; 95US-00484631.
PR 07-JUN-1995; 95US-00484635.
XX
PA (JOHJ) JOHNSON & JOHNSON CORP.
PA (AFFY-) AFFYMAX TECHNOLOGIES NV.
XX
PI Wrighton NC, Dower WJ, Chang RS, Kashyap AK, Jolliffe LK;
PI Johnson D, Mulcahy L;
XX
DR WPI; 1997-052225/05.
XX
PT Erythropoietin receptor binding peptide - useful for treating disorders
PT characterised by deficiency of EPO, or low or defective red blood cell
PT population.
XX
PS Disclosure; Fig 2; 95pp; English.
XX
CC The invention describes a peptide of 10-40 amino acid residues which
CC binds to erythropoietin (EPO) receptor and which includes the amino acid
CC sequence Cys-Xaa1-Xaa2-Gly-Pro-Xaa3-Thr-Trp-Xaa4-Cys, where Xaa1 = Arg,
CC His, Leu or Trp, Xaa2 = Met, Phe or Ile, Xaa3 = 1 of the 20 genetically
CC coded L-amino acids, and Xaa4 = Asp, Glu, Ile, Leu or Val. Optionally,
CC the peptide may be cyclised or dimerised. The peptide can be used to
CC treat a patient having a disorder characterised by a deficiency of EPO or
CC a low or defective red blood cell population. It can be used to treat end
CC stage renal failure or dialysis; anaemia associated with AIDS, autoimmune
CC disease, chronic inflammatory diseases or malignancy; beta-thalassaemia;
CC cystic fibrosis; early anaemia of prematurity; spinal cord injury; acute
CC blood loss; aging; and neoplastic disease states accompanied by abnormal
CC erythropoiesis. The peptides can also be used as reagents for detecting
CC EPO receptors on living cells, in biological fluids, in tissue
CC homogenates, etc. Sequences AAY13662-735 are representative peptides of
CC the invention
XX
SQ Sequence 22 AA;

```
  Query Match             94.4%;  Score 51;  DB 2;  Length 22;
  Best Local Similarity   58.3%;  Pred. No. 0.15;
  Matches    7;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 YXCXXGPXTWXC 13
              | |  || || |
Db          4 YSCFMGPTTWVC 15
```

RESULT 10
AAY26491
ID AAY26491 standard; peptide; 22 AA.
XX
AC AAY26491;
XX
DT 06-SEP-1999 (first entry)
XX
DE Erythropoietin receptor (EPO-R) binding peptide.
XX
KW Erythropoietin; EPO; receptor; EPO deficiency; renal failure; AIDS;
KW dialysis; anaemia; autoimmune disease; chronic inflammatory disease;
KW malignancy; beta-thalassaemia; cystic fibrosis; prematurity; blood loss;
KW spinal cord injury; aging; neoplastic disease; erythropoiesis; EPO-R.
XX
OS Synthetic.
XX

PN WO9640749-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US009810.
XX
PR 07-JUN-1995; 95US-00484631.
PR 07-JUN-1995; 95US-00484635.
XX
PA (JOHJ) JOHNSON & JOHNSON CORP.
PA (AFFY-) AFFYMAX TECHNOLOGIES NV.
XX
PI Wrighton NC, Dower WJ, Chang RS, Kashyap AK, Jolliffe LK;
PI Johnson D, Mulcahy L;
XX
DR WPI; 1997-052225/05.
XX
PT Erythropoietin receptor binding peptide - useful for treating disorders
PT characterised by deficiency of EPO, or low or defective red blood cell
PT population.
XX
PS Disclosure; Page 23; 95pp; English.
XX
CC The invention describes a peptide of 10-40 amino acid residues which
CC binds to erythropoietin (EPO) receptor and which includes the amino acid
CC sequence Cys-Xaa1-Xaa2-Gly-Pro-Xaa3-Thr-Trp-Xaa4-Cys, where Xaa1 = Arg,
CC His, Leu or Trp, Xaa2 = Met, Phe or Ile, Xaa3 = 1 of the 20 genetically
CC coded L-amino acids, and Xaa4 = Asp, Glu, Ile, Leu or Val. Optionally,
CC the peptide may be cyclised or dimerised. The peptide can be used to
CC treat a patient having a disorder characterised by a deficiency of EPO or
CC a low or defective red blood cell population. It can be used to treat end
CC stage renal failure or dialysis; anaemia associated with AIDS, autoimmune
CC disease, chronic inflammatory diseases or malignancy; beta-thalassaemia;
CC cystic fibrosis; early anaemia of prematurity; spinal cord injury; acute
CC blood loss; aging; and neoplastic disease states accompanied by abnormal
CC erythropoiesis. The peptides can also be used as reagents for detecting
CC EPO receptors on living cells, in biological fluids, in tissue
CC homogenates, etc. Sequences AAY26352-548 are representative peptides
CC falling within the above peptide motif and isolated during the affinity
CC selection process
XX
SQ Sequence 22 AA;

Query Match 94.4%; Score 51; DB 2; Length 22;
Best Local Similarity 58.3%; Pred. No. 0.15;
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 YXCXXGPXTWXC 13
|| || || | |
Db 5 YSCHMGPSTWVC 16

RESULT 11
AAY26355
ID AAY26355 standard; peptide; 22 AA.
XX
AC AAY26355;
XX
DT 06-SEP-1999 (first entry)
XX
DE Erythropoietin receptor (EPO-R) binding peptide.
XX
KW Erythropoietin; EPO; receptor; EPO deficiency; renal failure; AIDS;
KW dialysis; anaemia; autoimmune disease; chronic inflammatory disease;
KW malignancy; beta-thalassaemia; cystic fibrosis; prematurity; blood loss;
KW spinal cord injury; aging; neoplastic disease; erythropoiesis; EPO-R.
XX
OS Synthetic.
XX
PN WO9640749-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US009810.
XX
PR 07-JUN-1995; 95US-00484631.
PR 07-JUN-1995; 95US-00484635.
XX
PA (JOHJ) JOHNSON & JOHNSON CORP.
PA (AFFY-) AFFYMAX TECHNOLOGIES NV.
XX
PI Wrighton NC, Dower WJ, Chang RS, Kashyap AK, Jolliffe LK;
PI Johnson D, Mulcahy L;
XX
DR WPI; 1997-052225/05.
XX
PT Erythropoietin receptor binding peptide - useful for treating disorders
PT characterised by deficiency of EPO, or low or defective red blood cell
PT population.
XX
PS Disclosure; Page 16; 95pp; English.
XX
CC The invention describes a peptide of 10-40 amino acid residues which
CC binds to erythropoietin (EPO) receptor and which includes the amino acid
CC sequence Cys-Xaa1-Xaa2-Gly-Pro-Xaa3-Thr-Trp-Xaa4-Cys, where Xaa1 = Arg,
CC His, Leu or Trp, Xaa2 = Met, Phe or Ile, Xaa3 = 1 of the 20 genetically
CC coded L-amino acids, and Xaa4 = Asp, Glu, Ile, Leu or Val. Optionally,
CC the peptide may be cyclised or dimerised. The peptide can be used to
CC treat a patient having a disorder characterised by a deficiency of EPO or
CC a low or defective red blood cell population. It can be used to treat end
CC stage renal failure or dialysis; anaemia associated with AIDS, autoimmune
CC disease, chronic inflammatory diseases or malignancy; beta-thalassaemia;
CC cystic fibrosis; early anaemia of prematurity; spinal cord injury; acute
CC blood loss; aging; and neoplastic disease states accompanied by abnormal
CC erythropoiesis. The peptides can also be used as reagents for detecting
CC EPO receptors on living cells, in biological fluids, in tissue
CC homogenates, etc. Sequences AAY26352-548 are representative peptides
CC falling within the above peptide motif and isolated during the affinity
CC selection process
XX
SQ Sequence 22 AA;

Query Match 94.4%; Score 51; DB 2; Length 22;
Best Local Similarity 58.3%; Pred. No. 0.15;
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 YXCXXGPXTWXC 13
|| || || | |
Db 5 YSCHMGPSTWVC 16

RESULT 12
AAW27023
ID AAW27023 standard; peptide; 22 AA.
XX
AC AAW27023;
XX
DT 11-NOV-1997 (first entry)
XX
DE Monomer subunit of erythropoietin receptor binding dimer.
XX
KW Monomer; erythropoietin; EPO; receptor; binding; dimer; activation;
KW treatment; disorder; deficiency; low; defective; red blood cell;
KW erythrocyte; population; cell surface; agonist; end stage; renal;
KW failure; dialysis; anaemia; anemia; AIDS; chronic; inflammatory; disease;
KW rheumatoid arthritis; bowel inflammation; autoimmune; transfusion.
XX
OS Synthetic.
XX
PN WO9640772-A2.
XX
PD 19-DEC-1996.
XX
PF 06-JUN-1996; 96WO-US009469.
XX
PR 07-JUN-1995; 95US-00484135.

```
XX
PA   (JOHJ ) JOHNSON & JOHNSON.
XX
PI   Johnson DL,  Zivin RA;
XX
DR   WPI; 1997-099920/09.
XX
PT   Activating cell surface receptors using peptide dimer agonists - also,
PT   new dimers of erythropoietin receptor binding peptide(s) useful for
PT   treating patient having disorder characterised by EPO deficiency.
XX
PS   Disclosure; Fig 9; 110pp; English.
XX
CC   The present peptide is a specific example of a claimed generic monomer
CC   subunit of an erythropoietin (EPO) receptor binding dimer, which
CC   comprises 2 EPO receptor binding monomers of 10 to 40 amino acids, and
CC   activates or improves the bioactivity of the EPO cell surface receptor.
CC   The dimer can be used to treat disorders resulting from EPO deficiency by
CC   improving the activity of its cell surface receptor, e.g. end stage renal
CC   failure/dialysis, anaemia associated with AIDS or chronic inflammatory
CC   diseases such as rheumatoid arthritis and chronic bowel inflammation and
CC   autoimmune disease. It can also be used to boost the red cell count of a
CC   patient prior to surgery or as pretreatment to transfusion. The dimer
CC   peptide exhibits increased biological potency in vitro and in vivo
CC   relative to its component monomeric agonists. Dimerisation may also
CC   convert cell surface receptor antagonists into agonists
XX
SQ   Sequence 22 AA;

  Query Match             94.4%;  Score 51;  DB 2;  Length 22;
  Best Local Similarity   58.3%;  Pred. No. 0.15;
  Matches    7;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 YXCXXGPXTWXC 13
              | |  || || |
Db          4 YSCFMGPTTWVC 15


RESULT 13
ADU91911
ID   ADU91911 standard; peptide; 22 AA.
XX
AC   ADU91911;
XX
DT   10-FEB-2005  (first entry)
XX
DE   EPO-R agonist SEQ ID NO 52.
XX
KW   erythropoietin receptor; EPO-R; erythropoietin; renal failure;
KW   autoimmune disease; cystic fibrosis; anemia; inflammation;
KW   spinal cord injury; aging; neurological disease; nephrotropic;
KW   antianemic; immunosuppressive; CNS-Gen.; neuroprotective;
KW   respiratory-Gen.; antiinflammatory; vulnerary; nootropic; cytostatic;
KW   hemostatic; cyclic.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "Acetylated residue"
FT   Disulfide-bond  7. .16
FT   Modified-site   22
FT                   /note= "C-terminal amide"
XX
PN   WO2004101611-A2.
XX
PD   25-NOV-2004.
XX
PF   12-MAY-2004; 2004WO-US014886.
XX
PR   12-MAY-2003; 2003US-0470245P.
XX
PA   (AFFY-) AFFYMAX INC.
XX
PI   Yin K,  Holmes C,  Lalonde G,  Balu P,  Schatz PJ,  Tumelty D;
XX
DR   WPI; 2005-039329/04.
XX
PT   New peptide comprising specified sequence of amino acid is erythropoietin
PT   receptor agonist useful for treating e.g. anemia, beta-thalassemia, renal
PT   disorders.
XX
PS   Disclosure; SEQ ID NO 52; 83pp; English.
XX
CC   This invention describes a novel peptide which is an erythropoietin
CC   receptor (EPO-R) activator. The peptide forms a dimer comprising a
CC   linking moiety connecting two peptide chains composed of ADU91861. The N-
CC   terminal of the peptide is acetylated. The EPO-R activator further
CC   comprises at least one water soluble polymer, preferably polyethylene
CC   glycol (PEG) covalently bound to the peptide and a spacer moiety. The
CC   products of the invention are used for treating disorders associated with
CC   deficiency of erythropoietin or low or defective red blood cell
CC   population, end stage renal failure or dialysis, anemia associated with
CC   AIDS, autoimmune disease or malignancy, beta-thalassemia, cystic
CC   fibrosis, early anemia of prematurity, anemia associated with chronic
CC   inflammatory disease, spinal cord injury, acute blood loss, aging and
CC   neoplastic disease states accompanied by abnormal erythropoiesis. The
CC   peptide compounds are potent agonists of erythropoietin receptor and have
CC   nephrotropic, antianemic, immunosuppressive, CNS-Gen., neuroprotective,
CC   respiratory-Gen., antiinflammatory, vulnerary, nootropic, cytostatic and
CC   hemostatic activity. This sequence represents a peptide which acts as an
CC   erythropoietin receptor (EPO-R) agonist.
XX
SQ   Sequence 22 AA;

  Query Match             94.4%;  Score 51;  DB 9;  Length 22;
  Best Local Similarity   58.3%;  Pred. No. 0.15;
  Matches    7;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 YXCXXGPXTWXC 13
              | |  || || |
Db          5 YSCHMGPSTWVC 16


RESULT 14
ADU91986
ID   ADU91986 standard; peptide; 23 AA.
XX
AC   ADU91986;
XX
DT   10-FEB-2005  (first entry)
XX
DE   EPO-R agonist SEQ ID NO 127.
XX
KW   erythropoietin receptor; EPO-R; erythropoietin; renal failure;
KW   autoimmune disease; cystic fibrosis; anemia; inflammation;
KW   spinal cord injury; aging; neurological disease; nephrotropic;
KW   antianemic; immunosuppressive; CNS-Gen.; neuroprotective;
KW   respiratory-Gen.; antiinflammatory; vulnerary; nootropic; cytostatic;
KW   hemostatic; cyclic.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "Acetylated residue"
FT   Disulfide-bond  7. .16
FT   Modified-site   23
FT                   /note= "C-terminal amide"
XX
PN   WO2004101611-A2.
XX
PD   25-NOV-2004.
XX
PF   12-MAY-2004; 2004WO-US014886.
XX
```

```
PR   12-MAY-2003; 2003US-0470245P.
XX
PA   (AFFY-) AFFYMAX INC.
XX
PI   Yin K,  Holmes C,  Lalonde G,  Balu P,  Schatz PJ,  Tumelty D;
XX
DR   WPI; 2005-039329/04.
XX
PT   New peptide comprising specified sequence of amino acid is erythropoietin
PT   receptor agonist useful for treating e.g. anemia, beta-thalassemia, renal
PT   disorders.
XX
PS   Disclosure; SEQ ID NO 127; 83pp; English.
XX
CC   This invention describes a novel peptide which is an erythropoietin
CC   receptor (EPO-R) activator. The peptide forms a dimer comprising a
CC   linking moiety connecting two peptide chains composed of ADU91861. The N-
CC   terminal of the peptide is acetylated. The EPO-R activator further
CC   comprises at least one water soluble polymer, preferably polyethylene
CC   glycol (PEG) covalently bound to the peptide and a spacer moiety. The
CC   products of the invention are used for treating disorders associated with
CC   deficiency of erythropoietin or low or defective red blood cell
CC   population, end stage renal failure or dialysis, anemia associated with
CC   AIDS, autoimmune disease or malignancy, beta-thalassemia, cystic
CC   fibrosis, early anemia of prematurity, anemia associated with chronic
CC   inflammatory disease, spinal cord injury, acute blood loss, aging and
CC   neoplastic disease states accompanied by abnormal erythropoiesis. The
CC   peptide compounds are potent agonists of erythropoietin receptor and have
CC   nephrotropic, antianemic, immunosuppressive, CNS-Gen., neuroprotective,
CC   respiratory-Gen., antiinflammatory, vulnerary, nootropic, cytostatic and
CC   hemostatic activity. This sequence represents a peptide which acts as an
CC   erythropoietin receptor (EPO-R) agonist.
XX
SQ   Sequence 23 AA;

  Query Match             94.4%;  Score 51;  DB 9;  Length 23;
  Best Local Similarity   58.3%;  Pred. No. 0.16;
  Matches    7;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 YXCXXGPXTWXC 13
              | |  || || |
Db          5 YSCTMGPRTWVC 16


RESULT 15
ADU91945
ID   ADU91945 standard; peptide; 23 AA.
XX
AC   ADU91945;
XX
DT   10-FEB-2005  (first entry)
XX
DE   EPO-R agonist SEQ ID NO 86.
XX
KW   erythropoietin receptor; EPO-R; erythropoietin; renal failure;
KW   autoimmune disease; cystic fibrosis; anemia; inflammation;
KW   spinal cord injury; aging; neurological disease; nephrotropic;
KW   antianemic; immunosuppressive; CNS-Gen.; neuroprotective;
KW   respiratory-Gen.; antiinflammatory; vulnerary; nootropic; cytostatic;
KW   hemostatic; cyclic.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "Acetylated residue"
FT   Disulfide-bond  7. .16
FT   Modified-site   23
FT                   /note= "C-terminal amide"
XX
PN   WO2004101611-A2.
XX
PD   25-NOV-2004.
XX
PF   12-MAY-2004; 2004WO-US014886.
XX
PR   12-MAY-2003; 2003US-0470245P.
XX
PA   (AFFY-) AFFYMAX INC.
XX
PI   Yin K,  Holmes C,  Lalonde G,  Balu P,  Schatz PJ,  Tumelty D;
XX
DR   WPI; 2005-039329/04.
XX
PT   New peptide comprising specified sequence of amino acid is erythropoietin
PT   receptor agonist useful for treating e.g. anemia, beta-thalassemia, renal
PT   disorders.
XX
PS   Disclosure; SEQ ID NO 86; 83pp; English.
XX
CC   This invention describes a novel peptide which is an erythropoietin
CC   receptor (EPO-R) activator. The peptide forms a dimer comprising a
CC   linking moiety connecting two peptide chains composed of ADU91861. The N-
CC   terminal of the peptide is acetylated. The EPO-R activator further
CC   comprises at least one water soluble polymer, preferably polyethylene
CC   glycol (PEG) covalently bound to the peptide and a spacer moiety. The
CC   products of the invention are used for treating disorders associated with
CC   deficiency of erythropoietin or low or defective red blood cell
CC   population, end stage renal failure or dialysis, anemia associated with
CC   AIDS, autoimmune disease or malignancy, beta-thalassemia, cystic
CC   fibrosis, early anemia of prematurity, anemia associated with chronic
CC   inflammatory disease, spinal cord injury, acute blood loss, aging and
CC   neoplastic disease states accompanied by abnormal erythropoiesis. The
CC   peptide compounds are potent agonists of erythropoietin receptor and have
CC   nephrotropic, antianemic, immunosuppressive, CNS-Gen., neuroprotective,
CC   respiratory-Gen., antiinflammatory, vulnerary, nootropic, cytostatic and
CC   hemostatic activity. This sequence represents a peptide which acts as an
CC   erythropoietin receptor (EPO-R) agonist.
XX
SQ   Sequence 23 AA;

  Query Match             94.4%;  Score 51;  DB 9;  Length 23;
  Best Local Similarity   58.3%;  Pred. No. 0.16;
  Matches    7;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 YXCXXGPXTWXC 13
              | |  || || |
Db          5 YSCTMGPRTWVC 16


Search completed: March 31, 2006, 16:22:28
Job time : 62.6915 secs
```

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: March 31, 2006, 16:22:51 ; Search time 9.95025 Seconds
(without alignments)
154.717 Million cell updates/sec

Title: US-10-609-217-420
Perfect score: 54
Sequence: 1 XYXCXXGPXTWXCXXX 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	45	83.3	19	1	EWSMAN	ancovenin - Strept
2	37	68.5	460	2	S06022	regulatory protein
3	37	68.5	475	2	H84137	hypothetical prote
4	36	66.7	123	2	I52427	guanine-nucleotide
5	36	66.7	123	2	S29714	guanine-nucleotide
6	36	66.7	571	1	S30253	GABA transport pro
7	35.5	65.7	4543	1	A53102	alpha-2-macroglobu
8	35	64.8	318	2	E87929	protein T22H2.6 [i
9	35	64.8	345	2	T25138	hypothetical prote
10	35	64.8	358	2	T25137	hypothetical prote
11	35	64.8	645	2	T27186	hypothetical prote
12	35	64.8	2531	2	S18188	notch protein homo
13	35	64.8	2531	2	A46019	notch-1 protein -
14	35	64.8	2555	2	A40043	notch protein homo
15	34.5	63.9	1661	2	T31330	head-activator bin
16	34	63.0	19	1	EWSMCN	cinnamycin - Strep
17	34	63.0	78	1	EWSMYG	cinnamycin precurs
18	34	63.0	1693	2	S76086	beta transducin-li
19	33.5	62.0	4544	1	S02392	alpha-2-macroglobu
20	33.5	62.0	4545	1	S25111	alpha-2-macroglobu
21	33	61.1	68	2	B43940	lactococcin B prec
22	33	61.1	217	2	H86188	protein T25N20.5 [
23	33	61.1	266	2	H86407	F3H9.15 protein -
24	33	61.1	292	2	G88071	protein ZK1240.5 [
25	33	61.1	410	2	S38238	hypothetical prote
26	33	61.1	449	2	AC0234	probable exported
27	33	61.1	449	2	T47039	hypothetical prote
28	33	61.1	500	2	T49388	related to ascus d
29	33	61.1	555	2	T38108	high-affinity gluc
30	33	61.1	557	2	T43657	probable glucose t
31	32	59.3	255	2	F39925	hypothetical prote
32	32	59.3	274	2	T10270	protein kinase (EC
33	32	59.3	279	2	G71429	hypothetical prote
34	32	59.3	292	2	S60997	ARG11 protein - ye
35	32	59.3	307	2	C81862	conserved hypothet
36	32	59.3	307	2	D81082	conserved hypothet
37	32	59.3	315	2	AD2298	hypothetical prote
38	32	59.3	540	2	S72233	transcription fact
39	32	59.3	568	2	JC5629	mullerian-inhibiti
40	32	59.3	704	2	F86146	hypothetical prote
41	32	59.3	733	2	A97415	hypothetical prote
42	32	59.3	840	2	T02164	hypothetical prote
43	32	59.3	1531	1	DVHUAR	multidrug resistan
44	31.5	58.3	1149	2	I38006	M130 antigen precu
45	31.5	58.3	1151	2	I38004	M130 antigen precu

ALIGNMENTS

RESULT 1
EWSMAN
ancovenin - Streptomyces sp. (strain A647P-2)
C;Species: Streptomyces sp.
C;Date: 12-May-1994 #sequence_revision 19-May-1994 #text_change 09-Jul-2004
C;Accession: A61284
R;Wakamiya, T.; Ueki, Y.; Shiba, T.; Kido, Y.; Motoki, Y.
Tetrahedron Lett. 26, 665-668, 1985
A;Title: The structure of ancovenin, a new peptide inhibitor of angiotensin I converting
A;Reference number: A61284
A;Accession: A61284
A;Molecule type: protein
A;Residues: 1-19 <WAK>
A;Cross-references: UNIPROT:P38655; UNIPARC:UPI0000052CC3
C;Superfamily: cinnamycin precursor
C;Keywords: antibiotic; lanthionine
F;1-18/Cross-link: (2S,3S,6R)-3-methyl-lanthionine (Cys-Thr) #status experimental
F;4-14/Cross-link: sn-(2S,6R)-lanthionine (Ser-Cys) #status experimental
F;5-11/Cross-link: (2S,3S,6R)-3-methyl-lanthionine (Cys-Thr) #status experimental
F;6/Modified site: dehydroalanine (Ser) #status experimental

Query Match 83.3%; Score 45; DB 1; Length 19;
Best Local Similarity 60.0%; Pred. No. 0.073;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy          4 CXXGPXTWXC 13
              |  ||  ||  |
Db          5 CSFGPLTWSC 14
```

RESULT 2
S06022
regulatory protein O2 - maize
C;Species: Zea mays (maize)
C;Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 31-Dec-2004
C;Accession: S06022; S06009
R;Hartings, H.; Maddaloni, M.; Lazzaroni, N.; di Fonzo, N.; Motto, M.; Salamini, F.; Thc
EMBO J. 8, 2795-2801, 1989
A;Title: The O2 gene which regulates zein deposition in maize endosperm encodes a protei
A;Reference number: S06022; MUID:90059860; PMID:2479535
A;Accession: S06022
A;Molecule type: mRNA
A;Residues: 1-460 <HAR>
A;Cross-references: UNIPROT:P12959; UNIPARC:UPI000016E05D; GB:X16618; NID:g22383; PIDN:C
R;Maddaloni, M.; di Fonzo, N.; Hartings, H.; Lazzaroni, N.; Salamini, F.; Thompson, R.;
Nucleic Acids Res. 17, 7532, 1989
A;Title: The sequence of the zein regulatory gene opaque-2 (02) of Zea Mays.
A;Reference number: S06009; MUID:90016825; PMID:2798113
A;Accession: S06009
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-22,29-149,'D',151-460 <MAD>

```
A;Cross-references: UNIPARC:UPI00001794F4; EMBL:X15544
C;Genetics:
A;Gene: opaque 2
A;Map position: 7
A;Introns: 148/3; 168/3; 238/2; 263/3; 305/3
C;Superfamily: BZIP protein; fos/jun DNA-binding domain homology
C;Keywords: DNA binding; nucleus; transcription regulation
F;227-267/Domain: fos/jun DNA-binding domain homology <FJD>

  Query Match             68.5%;  Score 37;  DB 2;  Length 460;
  Best Local Similarity   71.4%;  Pred. No. 30;
  Matches    5;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          7 GPXTWXC 13
              || || |
Db        436 GPYTWTC 442


RESULT 3
H84137
hypothetical protein BH3904 [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C;Accession: H84137
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: H84137
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-475 <STO>
A;Cross-references: UNIPROT:Q9K628; UNIPARC:UPI00000C432F; GB:AP001520; GB:BA000004; NID
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH3904

  Query Match             68.5%;  Score 37;  DB 2;  Length 475;
  Best Local Similarity   62.5%;  Pred. No. 31;
  Matches    5;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          4 CXXGPXTW 11
              |  || ||
Db        156 CAGGPSTW 163


RESULT 4
I52427
guanine-nucleotide-releasing protein Mss4 - human
C;Species: Homo sapiens (man)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C;Accession: I52427
R;Yu, H.; Schreiber, S.L.
Biochemistry 34, 9103-9110, 1995
A;Title: Cloning, Zn2+ binding, and structural characterization of the guanine nucleotid
A;Reference number: I52427; MUID:95345082; PMID:7619808
A;Accession: I52427
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-123 <RES>
A;Cross-references: UNIPROT:P47224; UNIPARC:UPI00001117CC; GB:S78873; NID:g1037135; PIDN
C;Genetics:
A;Gene: GDB:MSS4
A;Cross-references: GDB:683578

  Query Match             66.7%;  Score 36;  DB 2;  Length 123;
  Best Local Similarity   50.0%;  Pred. No. 14;
  Matches    5;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          4 CXXGPXTWXC 13
              |  ||  | |
Db         97 CEIGPIGWHC 106


RESULT 5
S29714
guanine-nucleotide-releasing protein mss4 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C;Accession: S29714
R;Burton, J.; Roberts, D.; Montaldi, M.; Novick, P.; de Camilli, P.
Nature 361, 464-467, 1993
A;Title: A mammalian guanine-nucleotide-releasing protein enhances function of yeast sec
A;Reference number: S29714; MUID:93156814; PMID:8429887
A;Accession: S29714
A;Molecule type: mRNA
A;Residues: 1-123 <BUR>
A;Cross-references: UNIPROT:Q08326; UNIPARC:UPI000012F68D; EMBL:X70496; NID:g313871; PIDN
C;Genetics:
A;Gene: mss4

  Query Match             66.7%;  Score 36;  DB 2;  Length 123;
  Best Local Similarity   50.0%;  Pred. No. 14;
  Matches    5;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          4 CXXGPXTWXC 13
              |  ||  | |
Db         97 CEIGPIGWHC 106


RESULT 6
S30253
GABA transport protein - yeast (Saccharomyces cerevisiae)
N;Alternate names: GABA-specific permease; protein D1037; protein YDL210w
C;Species: Saccharomyces cerevisiae
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: S30253; S67769; S25147
R;Andre, B.; Hein, C.; Grenson, M.; Jauniaux, J.C.
Mol. Gen. Genet. 237, 17-25, 1993
A;Title: Cloning and expression of the UGA4 gene coding for the inducible GABA-specific t
A;Reference number: S30253; MUID:93204891; PMID:8455553
A;Accession: S30253
A;Molecule type: DNA
A;Residues: 1-571 <AND>
A;Cross-references: UNIPROT:P32837; UNIPARC:UPI0000137AE9; EMBL:X66472; NID:g4749; PIDN:C
A;Note: the sequence from Fig. 5 is inconsistent with that from Fig. 3 in having 527-X
R;Schmidt, E.R.; Bahr, A.; Kraemer, C.; Hankeln, T.; Moeller-Rieker, S.
submitted to the Protein Sequence Database, July 1996
A;Reference number: S67756
A;Accession: S67769
A;Molecule type: DNA
A;Residues: 1-571 <SCH>
A;Cross-references: UNIPARC:UPI0000137AE9; EMBL:Z74258; NID:g1431349; PIDN:CAA98788.1; PI
A;Experimental source: strain S288C
C;Genetics:
A;Gene: SGD:UGA4; MIPS:YDL210w
A;Cross-references: SGD:S0002369; MIPS:YDL210w
A;Map position: 4L
C;Superfamily: choline transport protein
C;Keywords: transmembrane protein
F;79-95/Domain: transmembrane #status predicted <TM1>
F;108-124/Domain: transmembrane #status predicted <TM2>
F;154-170/Domain: transmembrane #status predicted <TM3>
F;203-219/Domain: transmembrane #status predicted <TM4>
F;229-245/Domain: transmembrane #status predicted <TM5>
F;320-336/Domain: transmembrane #status predicted <TM6>
F;365-381/Domain: transmembrane #status predicted <TM7>
F;420-436/Domain: transmembrane #status predicted <TM8>
F;485-501/Domain: transmembrane #status predicted <TM9>

  Query Match             66.7%;  Score 36;  DB 1;  Length 571;
  Best Local Similarity   50.0%;  Pred. No. 54;
  Matches    5;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 YXCXXGPXTW 11
              | |  ||  |
```

```
Db        517 YACVIGPGIW 526


RESULT 7
A53102
alpha-2-macroglobulin receptor precursor - chicken
N;Alternate names: CD91; LDL receptor-related protein 1; low density lipoprotein receptc
C;Species: Gallus gallus (chicken)
C;Date: 04-Sep-1998 #sequence_revision 04-Sep-1998 #text_change 09-Jul-2004
C;Accession: A53102
R;Nimpf, J.; Stifani, S.; Bilous, P.T.; Schneider, W.J.
J. Biol. Chem. 269, 212-219, 1994
A;Title: The somatic cell-specific low density lipoprotein receptor-related protein of t
A;Reference number: A53102; MUID:94103212; PMID:7506255
A;Accession: A53102
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-4543 <NIM>
A;Cross-references: UNIPROT:P98157; UNIPARC:UPI000013C4B3; GB:X74904; NID:g438006; PIDN:
C;Complex: The alpha-2-macroglobulin receptor complex consists of noncovalently-associat
d protein.
C;Superfamily: alpha-2-macroglobulin receptor; EGF homology; LDL receptor ligand-binding
C;Keywords: beta-hydroxyasparagine; beta-hydroxyaspartic acid; calcium binding; glycoprc
F;1-17/Domain: signal sequence #status predicted <SIG>
F;18-3942/Domain: alpha-2-macroglobulin receptor 515K chain #status predicted <515K>
F;18-3942,3943-4543/Product: alpha-2-macroglobulin receptor #status predicted <MAT>
F;29-66/Domain: LDL receptor ligand-binding repeat homology <LDL1>
F;74-110/Domain: LDL receptor ligand-binding repeat homology <LDL2>
F;117-150/Domain: EGF homology <EG1>
F;156-190/Domain: EGF homology <EG2>
F;200-241/Domain: LDL receptor YWTD-containing repeat homology <YW01>
F;242-283/Domain: LDL receptor YWTD-containing repeat homology <YW02>
F;294-336/Domain: LDL receptor YWTD-containing repeat homology <YW03>
F;337-380/Domain: LDL receptor YWTD-containing repeat homology <YW04>
F;381-422/Domain: LDL receptor YWTD-containing repeat homology <YW05>
F;423-470/Domain: LDL receptor YWTD-containing repeat homology <YW06>
F;480-521/Domain: EGF homology <EG3>
F;573-615/Domain: LDL receptor YWTD-containing repeat homology <YW07>
F;616-661/Domain: LDL receptor YWTD-containing repeat homology <YW08>
F;662-712/Domain: LDL receptor YWTD-containing repeat homology <YW09>
F;713-754/Domain: LDL receptor YWTD-containing repeat homology <YW10>
F;755-797/Domain: LDL receptor YWTD-containing repeat homology <YW11>
F;805-840/Domain: EGF homology <EG4>
F;852-888/Domain: LDL receptor ligand-binding repeat homology <LDL3>
F;893-929/Domain: LDL receptor ligand-binding repeat homology <LDL4>
F;934-969/Domain: LDL receptor ligand-binding repeat homology <LDL5>
F;974-1009/Domain: LDL receptor ligand-binding repeat homology <LDL6>
F;1013-1049/Domain: LDL receptor ligand-binding repeat homology <LDL7>
F;1060-1095/Domain: LDL receptor ligand-binding repeat homology <LDL8>
F;1102-1138/Domain: LDL receptor ligand-binding repeat homology <LDL9>
F;1143-1180/Domain: LDL receptor ligand-binding repeat homology <LDLA>
F;1183-1219/Domain: EGF homology <EG5>
F;1225-1259/Domain: EGF homology <EG6>
F;1267-1306/Domain: LDL receptor YWTD-containing repeat homology <YW12>
F;1307-1353/Domain: LDL receptor YWTD-containing repeat homology <YW13>
F;1354-1396/Domain: LDL receptor YWTD-containing repeat homology <YW14>
F;1397-1443/Domain: LDL receptor YWTD-containing repeat homology <YW15>
F;1444-1486/Domain: LDL receptor YWTD-containing repeat homology <YW16>
F;1487-1529/Domain: LDL receptor YWTD-containing repeat homology <YW17>
F;1538-1576/Domain: EGF homology <EG7>
F;1581-1624/Domain: LDL receptor YWTD-containing repeat homology <YW18>
F;1625-1667/Domain: LDL receptor YWTD-containing repeat homology <YW19>
F;1668-1711/Domain: LDL receptor YWTD-containing repeat homology <YW20>
F;1712-1751/Domain: LDL receptor YWTD-containing repeat homology <YW21>
F;1752-1794/Domain: LDL receptor YWTD-containing repeat homology <YW22>
F;1795-1842/Domain: LDL receptor YWTD-containing repeat homology <YW23>
F;1846-1882/Domain: EGF homology <EG8>
F;1930-1972/Domain: LDL receptor YWTD-containing repeat homology <YW24>
F;1973-2015/Domain: LDL receptor YWTD-containing repeat homology <YW25>
F;2016-2059/Domain: LDL receptor YWTD-containing repeat homology <YW26>
F;2060-2101/Domain: LDL receptor YWTD-containing repeat homology <YW27>
F;2102-2147/Domain: LDL receptor YWTD-containing repeat homology <YW28>
F;2155-2190/Domain: EGF homology <EG9>
F;2195-2237/Domain: LDL receptor YWTD-containing repeat homology <YW29>
F;2247-2288/Domain: LDL receptor YWTD-containing repeat homology <YW30>
F;2338-2382/Domain: LDL receptor YWTD-containing repeat homology <YW31>
F;2383-2423/Domain: LDL receptor YWTD-containing repeat homology <YW32>
F;2424-2467/Domain: LDL receptor YWTD-containing repeat homology <YW33>
F;2476-2511/Domain: EGF homology <EG10>
F;2518-2555/Domain: LDL receptor ligand-binding repeat homology <LDLB>
F;2560-2594/Domain: LDL receptor ligand-binding repeat homology <LDLC>
F;2599-2633/Domain: LDL receptor ligand-binding repeat homology <LDLD>
F;2646-2682/Domain: LDL receptor ligand-binding repeat homology <LDLE>
F;2690-2724/Domain: LDL receptor ligand-binding repeat homology <LDLF>
F;2732-2767/Domain: LDL receptor ligand-binding repeat homology <LDLG>
F;2772-2810/Domain: LDL receptor ligand-binding repeat homology <LDLH>
F;2816-2851/Domain: LDL receptor ligand-binding repeat homology <LDLI>
F;2856-2895/Domain: LDL receptor ligand-binding repeat homology <LDLJ>
F;2902-2936/Domain: LDL receptor ligand-binding repeat homology <LDLK>
F;2941-2977/Domain: EGF homology <EG11>
F;2983-3018/Domain: EGF homology <EG12>
F;3026-3065/Domain: LDL receptor YWTD-containing repeat homology <YW34>
F;3066-3110/Domain: LDL receptor YWTD-containing repeat homology <YW35>
F;3111-3153/Domain: LDL receptor YWTD-containing repeat homology <YW36>
F;3154-3197/Domain: LDL receptor YWTD-containing repeat homology <YW37>
F;3198-3238/Domain: LDL receptor YWTD-containing repeat homology <YW38>
F;3239-3281/Domain: LDL receptor YWTD-containing repeat homology <YW39>
F;3291-3327/Domain: EGF homology <EG13>
F;3331-3366/Domain: LDL receptor ligand-binding repeat homology <LDLL>
F;3371-3405/Domain: LDL receptor ligand-binding repeat homology <LDLM>
F;3410-3445/Domain: LDL receptor ligand-binding repeat homology <LDLN>
F;3450-3486/Domain: LDL receptor ligand-binding repeat homology <LDLO>
F;3491-3528/Domain: LDL receptor ligand-binding repeat homology <LDLP>
F;3533-3567/Domain: LDL receptor ligand-binding repeat homology <LDLQ>
F;3572-3606/Domain: LDL receptor ligand-binding repeat homology <LDLR>
F;3610-3644/Domain: LDL receptor ligand-binding repeat homology <LDLS>
F;3651-3687/Domain: LDL receptor ligand-binding repeat homology <LDLT>
F;3692-3728/Domain: LDL receptor ligand-binding repeat homology <LDLU>
F;3738-3774/Domain: LDL receptor ligand-binding repeat homology <LDLV>
F;3783-3820/Domain: EGF homology <EG14>
F;3826-3858/Domain: EGF homology <EG15>
F;3866-3909/Domain: LDL receptor YWTD-containing repeat homology <YW40>
F;3910-3968/Domain: LDL receptor YWTD-containing repeat homology <YW41>
F;3943-4543/Domain: alpha-2-macroglobulin receptor 85K chain #status predicted <85K>
F;3943-4420/Domain: 85K chain extracellular #status predicted <EXT>
F;3969-4011/Domain: LDL receptor YWTD-containing repeat homology <YW42>
F;4012-4055/Domain: LDL receptor YWTD-containing repeat homology <YW43>
F;4056-4098/Domain: LDL receptor YWTD-containing repeat homology <YW44>
F;4099-4141/Domain: LDL receptor YWTD-containing repeat homology <YW45>
F;4150-4181/Domain: EGF homology <EG16>
F;4199-4230/Domain: EGF homology <EG17>
F;4235-4266/Domain: EGF homology <EG18>
F;4271-4302/Domain: EGF homology <EG19>
F;4307-4338/Domain: EGF homology <EG20>
F;4343-4373/Domain: EGF homology <EG21>
F;4376-4408/Domain: EGF homology <EG22>
F;4421-4443/Domain: transmembrane #status predicted <TMM>
F;4444-4543/Domain: intracellular #status predicted <INT>
F;116,138,187,276,359,448,731,926,1048,1152,1153,1193,1216,1305,1509,1556,1573,1614,1643
3485,3659,3786,3837,3952,4074,4124,4178,4278/Binding site: carbohydrate (Asn) (covalent)
F;168,2995/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted
F;2955/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status predicted

  Query Match             65.7%;  Score 35.5;  DB 1;  Length 4543;
  Best Local Similarity   40.0%;  Pred. No. 4e+02;
  Matches    6;  Conservative    1;  Mismatches    5;  Indels    3;  Gaps    1;

Qy          2 YXCXXG---PXTWXC 13
              : |  |   | || |
Db       2695 FACPSGRCIPMTWTC 2709


RESULT 8
E87929
protein T22H2.6 [imported] - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
```

```
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Dec-2002
C;Accession: E87929
R;anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
A;Reference number: A75000; MUID:99069613; PMID:9851916
A;Note: see websites genome.wustl.edu/gsc/C_elegans/ and www_sanger.ac.uk/Projects/C_ele
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A;Accession: E87929
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-318 <STO>
A;Cross-references: UNIPARC:UPI0000177C8F; GB:chr_I; PIDN:CAB04752.1; PID:g3880056; GSPD
C;Genetics:
A;Gene: T22H2.6
A;Map position: 1
C;Superfamily: protein T22H2.6

  Query Match             64.8%;  Score 35;  DB 2;  Length 318;
  Best Local Similarity   50.0%;  Pred. No. 49;
  Matches    5;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          4 CXXGPXTWXC 13
              |  |  || |
Db         71 CKLGDNTWGC 80


RESULT 9
T25138
hypothetical protein T22H2.6b - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T25138
R;Lennard, N.
submitted to the EMBL Data Library, November 1996
A;Reference number: Z19985
A;Accession: T25138
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-345 <WIL>
A;Cross-references: UNIPROT:Q9U362; UNIPARC:UPI000002A1D2; EMBL:Z81595; PIDN:CAB54305.1;
A;Experimental source: clone T22H2
C;Genetics:
A;Gene: CESP:T22H2.6b
A;Map position: 1
A;Introns: 93/3; 232/3; 314/3
C;Superfamily: protein T22H2.6

  Query Match             64.8%;  Score 35;  DB 2;  Length 345;
  Best Local Similarity   50.0%;  Pred. No. 53;
  Matches    5;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          4 CXXGPXTWXC 13
              |  |  || |
Db        111 CKLGDNTWGC 120


RESULT 10
T25137
hypothetical protein T22H2.6a - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T25137
R;Lennard, N.
submitted to the EMBL Data Library, November 1996
A;Reference number: Z19985
A;Accession: T25137
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-358 <WIL>
A;Cross-references: UNIPROT:Q9U362; UNIPARC:UPI000008667D; EMBL:Z81595; PIDN:CAB54304.1;
A;Experimental source: clone T22H2
C;Genetics:
A;Gene: CESP:T22H2.6a
A;Map position: 1
A;Introns: 93/3; 232/3; 314/3
C;Superfamily: protein T22H2.6

  Query Match             64.8%;  Score 35;  DB 2;  Length 358;
  Best Local Similarity   50.0%;  Pred. No. 55;
  Matches    5;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          4 CXXGPXTWXC 13
              |  |  || |
Db        111 CKLGDNTWGC 120


RESULT 11
T27186
hypothetical protein Y54G9A.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T27186
R;Smye, R.
submitted to the EMBL Data Library, October 1998
A;Reference number: Z20324
A;Accession: T27186
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-645 <WIL>
A;Cross-references: UNIPROT:Q9XWG9; UNIPARC:UPI000016404E; EMBL:AL032648; PIDN:CAA21699.1
A;Experimental source: clone Y54G9A
C;Genetics:
A;Gene: CESP:Y54G9A.3
A;Map position: 2
A;Introns: 56/3; 100/2; 148/1; 411/2; 541/2; 575/3

  Query Match             64.8%;  Score 35;  DB 2;  Length 645;
  Best Local Similarity   41.7%;  Pred. No. 91;
  Matches    5;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          2 YXCXXGPXTWXC 13
              | |   |  | | |
Db          8 YNCLERPTGWKC 19


RESULT 12
S18188
notch protein homolog - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 02-Aug-2002
C;Accession: S18188
R;Weinmaster, G.; Roberts, V.J.; Lemke, G.
Development 113, 199-205, 1991
A;Title: A homolog of Drosophila Notch expressed during mammalian development.
A;Reference number: S18188; MUID:92111383; PMID:1764995
A;Accession: S18188
A;Molecule type: mRNA
A;Residues: 1-2531 <WEI>
A;Cross-references: UNIPARC:UPI0000177456; EMBL:X57405; NID:g57634; PID:g57635
C;Superfamily: notch protein; ankyrin repeat homology; EGF homology
F;987-1018/Domain: EGF homology <EGF1>
F;1025-1056/Domain: EGF homology <EGF>
F;1233-1264/Domain: EGF homology <EGF2>
F;1917-1949/Domain: ankyrin repeat homology <AN1>
F;1950-1982/Domain: ankyrin repeat homology <AN2>
F;1984-2016/Domain: ankyrin repeat homology <AN3>
F;2017-2049/Domain: ankyrin repeat homology <AN4>
F;2050-2082/Domain: ankyrin repeat homology <AN5>

  Query Match             64.8%;  Score 35;  DB 2;  Length 2531;
  Best Local Similarity   50.0%;  Pred. No. 3e+02;
  Matches    5;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          4 CXXGPXTWXC 13
              |  || |: |
```

```
Db        543 CLDGPNTYTC 552


RESULT 13
A46019
notch-1 protein - mouse
N;Alternate names: motch protein
C;Species: Mus musculus (house mouse)
C;Date: 22-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 05-Oct-2004
C;Accession: A46019; S25144; C49175; B46438; A46438; PH1569; S32109
R;del Amo, F.F.; Gendron-Maguire, M.; Swiatek, P.J.; Jenkins, N.A.; Copeland, N.G.; Grid
Genomics 15, 259-264, 1993
A;Title: Cloning, analysis, and chromosomal localization of Notch-1, a mouse homolog of
A;Reference number: A46019; MUID:93194170; PMID:8449489
A;Accession: A46019
A;Status: not compared with conceptual translation
A;Molecule type: nucleic acid
A;Residues: 1-2531 <DEL>
A;Cross-references: UNIPROT:Q01705; UNIPARC:UPI000002922B; GB:Z11886; GB:S47228; NID:g28
A;Note: sequence extracted from NCBI backbone (NCBIP:127318)
R;Franco del Amo, F.; Smith, D.E.; Swiatek, P.J.; Gendron-Maguire, M.; Greenspan, R.J.;
submitted to the EMBL Data Library, April 1992
A;Description: Expression pattern of Motch, a mouse homolog of Drosophila Notch, suggest
A;Reference number: S25144
A;Accession: S25144
A;Molecule type: mRNA
A;Residues: 1551-2108,'Q',2110-2114,'ALP',2118-2170 <FRA>
A;Cross-references: UNIPARC:UPI0000177461; EMBL:Z11886
R;Lardelli, M.; Lendahl, U.
Exp. Cell Res. 204, 364-372, 1993
A;Title: Motch A and Motch B--two mouse Notch homologues coexpressed in a wide variety c
A;Reference number: A49175; MUID:93178563; PMID:8440332
A;Accession: C49175
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1161-1547 <LAR>
A;Cross-references: UNIPARC:UPI0000177462; EMBL:X68278; NID:g287987; PIDN:CAA48339.1; PI
A;Experimental source: embryo
A;Note: sequence extracted from NCBI backbone (NCBIP:126159)
R;Kopan, R.; Weintraub, H.
J. Cell Biol. 121, 631-641, 1993
A;Title: Mouse notch: expression in hair follicles correlates with cell fate determinati
A;Reference number: A46438; MUID:93252998; PMID:8486742
A;Accession: B46438
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1865-1932,'RR',1935-1937,'L',1938-1967,'I',1969-2044,'IE',2047-2052,'S',2054
A;Cross-references: UNIPARC:UPI0000177463
A;Experimental source: embryo
A;Note: sequence extracted from NCBI backbone (NCBIN:131246, NCBIP:131247)
C;Comment: This protein has many EGF repeats and lin-12[#1172]/Notch repeats.
C;Comment: This protein is one of the neurogenic proteins controlling the decision betwe
C;Genetics:
A;Gene: notch-1
A;Map position: 2
A;Note: proximal region of chromosome 2
C;Superfamily: notch protein; ankyrin repeat homology; EGF homology
F;106-138/Domain: EGF homology <EGF1>
F;144-175/Domain: EGF homology <EG01>
F;222-254/Domain: EGF homology <EGF2>
F;261-292/Domain: EGF homology <EG02>
F;339-370/Domain: EGF homology <EG03>
F;416-449/Domain: EGF homology <EGF3>
F;456-487/Domain: EGF homology <EG04>
F;494-525/Domain: EGF homology <EG05>
F;532-563/Domain: EGF homology <EG06>
F;607-638/Domain: EGF homology <EG07>
F;682-713/Domain: EGF homology <EG08>
F;757-788/Domain: EGF homology <EG09>
F;795-826/Domain: EGF homology <EG10>
F;873-904/Domain: EGF homology <EG11>
F;911-942/Domain: EGF homology <EG12>
F;949-980/Domain: EGF homology <EG13>
F;987-1018/Domain: EGF homology <EG14>
F;1025-1056/Domain: EGF homology <EG15>
F;1063-1094/Domain: EGF homology <EG16>
F;1149-1180/Domain: EGF homology <EG17>
F;1187-1218/Domain: EGF homology <EG18>
F;1233-1264/Domain: EGF homology <EGF4>
F;1352-1383/Domain: EGF homology <EG19>
F;1391-1425/Domain: EGF homology <EGF>
F;1917-1948/Domain: ankyrin repeat homology <AN1>
F;1949-1981/Domain: ankyrin repeat homology <AN2>
F;1983-2015/Domain: ankyrin repeat homology <AN3>
F;2016-2048/Domain: ankyrin repeat homology <AN4>
F;2049-2081/Domain: ankyrin repeat homology <AN5>

  Query Match             64.8%;  Score 35;  DB 2;  Length 2531;
  Best Local Similarity   50.0%;  Pred. No. 3e+02;
  Matches    5;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          4 CXXGPXTWXC 13
              |  ||  |: |
Db        543 CLDGPNTYTC 552


RESULT 14
A40043
notch protein homolog TAN-1 precursor - human
C;Species: Homo sapiens (man)
C;Date: 21-Apr-1992 #sequence_revision 21-Apr-1992 #text_change 05-Oct-2004
C;Accession: A40043
R;Ellisen, L.W.; Bird, J.; West, D.C.; Soreng, A.L.; Reynolds, T.C.; Smith, S.D.; Sklar,
Cell 66, 649-661, 1991
A;Title: TAN-1, the human homolog of the Drosophila Notch gene, is broken by chromosomal
A;Reference number: A40043; MUID:91347367; PMID:1831692
A;Accession: A40043
A;Status: preliminary; nucleic acid sequence not shown; not compared with conceptual trar
A;Molecule type: mRNA
A;Residues: 1-2555 <ELL>
A;Cross-references: UNIPARC:UPI0000177455; GB:M73980
C;Superfamily: notch protein; ankyrin repeat homology; EGF homology
F;261-292/Domain: EGF homology <EGX1>
F;494-525/Domain: EGF homology <EGF1>
F;987-1018/Domain: EGF homology <EGX2>
F;1149-1180/Domain: EGF homology <EGF>
F;1187-1218/Domain: EGF homology <EGF3>
F;1233-1264/Domain: EGF homology <EGX3>
F;1927-1959/Domain: ankyrin repeat homology <AN1>
F;1960-1992/Domain: ankyrin repeat homology <AN2>
F;1994-2026/Domain: ankyrin repeat homology <AN3>
F;2027-2059/Domain: ankyrin repeat homology <AN4>
F;2060-2092/Domain: ankyrin repeat homology <AN5>

  Query Match             64.8%;  Score 35;  DB 2;  Length 2555;
  Best Local Similarity   50.0%;  Pred. No. 3e+02;
  Matches    5;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          4 CXXGPXTWXC 13
              |  ||  |: |
Db        542 CLDGPNTYTC 551


RESULT 15
T31330
head-activator binding protein precursor - Chlorohydra viridissima
C;Species: Chlorohydra viridissima
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004
C;Accession: T31330
R;Hampe, W.; Franke, I.; Urny, J.; Petersen, C.M.; Schaller, H.C.
submitted to the EMBL Data Library, September 1998
A;Description: The neuropeptide head-activator binds to a new member of the low density
A;Reference number: Z20997
A;Accession: T31330
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
```

A;Residues: 1-1661 <HAM>
A;Cross-references: UNIPROT:O77244; UNIPARC:UPI000007E593; EMBL:AF092920; NID:g3719422;
C;Genetics:
A;Note: HAB

Query Match 63.9%; Score 34.5; DB 2; Length 1661;
Best Local Similarity 40.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 1; Mismatches 5; Indels 3; Gaps 1;

```
Qy          2 YXCXXG---PXTWXC 13
              : |  |   | || |
Db       1065 FKCTNGDCIPLTWKC 1079
```

Search completed: March 31, 2006, 16:37:21
Job time : 9.95025 secs

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: March 31, 2006, 16:09:36 ; Search time 59.9403 Seconds
(without alignments)
188.328 Million cell updates/sec

Title: US-10-609-217-420
Perfect score: 54
Sequence: 1 XYXCXXGPXTWXCXXX 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	45	83.3	19	1	DURC_STRGP	P36503 streptomyce
2	45	83.3	19	1	LANC_STRS6	P38655 streptomyce
3	41	75.9	532	2	Q8WWV6_HUMAN	Q8wwv6 homo sapien
4	41	75.9	534	2	Q96SA2_HUMAN	Q96sa2 homo sapien
5	41	75.9	577	2	Q5REH9_PONPY	Q5reh9 pongo pygma
6	41	75.9	589	2	Q5R770_PONPY	Q5r770 pongo pygma
7	40	74.1	167	2	Q6ZW93_HUMAN	Q6zw93 homo sapien
8	40	74.1	173	2	Q5VHX3_EAV	Q5vhx3 equine arte
9	40	74.1	180	2	Q4I355_GIBZE	Q4i355 gibberella
10	40	74.1	352	2	Q4IMN3_GIBZE	Q4imn3 gibberella
11	40	74.1	2022	2	Q61J27_CAEBR	Q61j27 caenorhabdi
12	39	72.2	172	2	Q6ZT75_HUMAN	Q6zt75 homo sapien
13	39	72.2	172	2	Q6ZWC2_HUMAN	Q6zwc2 homo sapien
14	39	72.2	173	2	Q9WD22_EAV	Q9wd22 equine arte
15	38	70.4	61	2	O70227_RAT	O70227 rattus norv
16	38	70.4	414	2	Q4SAV9_TETNG	Q4sav9 tetraodon n
17	37.5	69.4	2465	2	Q4RXZ7_TETNG	Q4rxz7 tetraodon n
18	37	68.5	157	2	Q6NEH5_CORDI	Q6neh5 corynebacte
19	37	68.5	378	1	TAZ_DROME	Q9v6g5 drosophila
20	37	68.5	389	2	Q84UZ1_CHLRE	Q84uz1 chlamydomon
21	37	68.5	453	1	OP2_MAIZE	P12959 zea mays (m
22	37	68.5	475	2	Q9K628_BACHD	Q9k628 bacillus ha
23	37	68.5	556	2	Q84UZ4_CHLRE	Q84uz4 chlamydomon
24	37	68.5	664	2	Q7QX73_GIALA	Q7qx73 giardia lam
25	37	68.5	741	2	Q4QE93_LEIMA	Q4qe93 leishmania
26	37	68.5	775	2	Q4H6M9_9DEIO	Q4h6m9 deinococcus
27	37	68.5	1192	2	Q7D3A2_AGRT5	Q7d3a2 agrobacteri
28	36	66.7	123	1	MSS4_HUMAN	P47224 homo sapien
29	36	66.7	123	1	MSS4_MOUSE	Q91x96 mus musculu
30	36	66.7	123	1	MSS4_RAT	Q08326 rattus norv
31	36	66.7	123	2	Q53EV1_HUMAN	Q53ev1 homo sapien
32	36	66.7	132	2	Q6DGQ2_BRARE	Q6dgq2 brachydanio
33	36	66.7	189	2	Q7MYI5_PHOLL	Q7myi5 photorhabdu
34	36	66.7	202	2	Q5YYW8_NOCFA	Q5yyw8 nocardia fa
35	36	66.7	234	2	Q6IG67_DROME	Q6ig67 drosophila
36	36	66.7	352	2	Q7UGA4_RHOBA	Q7uga4 rhodopirell
37	36	66.7	544	2	Q4SD11_TETNG	Q4sd11 tetraodon n
38	36	66.7	571	1	UGA4_YEAST	P32837 saccharomyc
39	36	66.7	581	2	Q9LKH1_MESCR	Q9lkh1 mesembryant
40	36	66.7	600	2	Q6MLD7_BDEBA	Q6mld7 bdellovibri
41	36	66.7	887	2	Q726W3_DESVH	Q726w3 desulfovibr
42	36	66.7	1623	2	Q4SS52_TETNG	Q4ss52 tetraodon n
43	35.5	65.7	536	2	Q6DG59_BRARE	Q6dg59 brachydanio
44	35.5	65.7	2304	2	Q4RW33_TETNG	Q4rw33 tetraodon n
45	35.5	65.7	4543	1	LRP1_CHICK	P98157 gallus gall

ALIGNMENTS

```
RESULT 1
DURC_STRGP
ID   DURC_STRGP     STANDARD;      PRT;    19 AA.
AC   P36503;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Lantibiotic duramycin C.
OS   Streptomyces griseoluteus.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=29306;
RN   [1]
RP   PROTEIN SEQUENCE.
RC   STRAIN=R2107;
RX   MEDLINE=91107436; PubMed=2125590;
RA   Fredenhagen A., Fendrich G., Marki F., Marki W., Gruner J.,
RA   Raschdorf F., Peter H.H.;
RT   "Duramycins B and C, two new lanthionine containing antibiotics as
RT   inhibitors of phospholipase A2. Structural revision of duramycin and
RT   cinnamycin.";
RL   J. Antibiot. 43:1403-1412(1990).
RN   [2]
RP   STRUCTURE BY NMR.
RA   Zimmermann N., Freund S., Fredenhagen A., Jung G.;
RT   "Solution structure of the lantibiotics duramycin B and C.";
RL   (In) Schneider C.H., Eberles A.N. (eds.);
RL   Peptides 1992, pp.519-520, Escom Science Publishers, Leiden (1993).
RN   [3]
RP   STRUCTURE BY NMR.
RX   MEDLINE=93387292; PubMed=8375380;
RA   Zimmermann N., Freund S., Fredenhagen A., Jung G.;
RT   "Solution structures of the lantibiotics duramycin B and C.";
RL   Eur. J. Biochem. 216:419-428(1993).
CC   -!- FUNCTION: Acts as inhibitor of phospholipase A2.
CC   -!- PTM: Maturation of lantibiotics involves the enzymic conversion of
CC       Thr, and Ser into dehydrated AA and the formation of thioether
CC       bonds with cysteine or the formation of dialkylamine bonds with
CC       lysine. This is followed by membrane translocation and cleavage of
CC       the modified precursor.
CC   -!- SIMILARITY: Belongs to the type B lantibiotic family.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
KW   Antibiotic; Antimicrobial; Bacteriocin; Direct protein sequencing;
KW   Lantibiotic; Thioether bond.
FT   CROSSLNK      1     18       Beta-methyllanthionine (Cys-Thr).
FT   CROSSLNK      4     14       Lanthionine (Ser-Cys).
FT   CROSSLNK      5     11       Beta-methyllanthionine (Cys-Thr).
FT   CROSSLNK      6     19       Lysinoalanine (Ser-Lys).
```

```
SQ   SEQUENCE   19 AA;  2007 MW;  E2404ECE3F95286A CRC64;

  Query Match             83.3%;  Score 45;  DB 1;  Length 19;
  Best Local Similarity   60.0%;  Pred. No. 0.26;
  Matches    6;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          4 CXXGPXTWXC 13
              |  ||  || |
Db          5 CSYGPLTWSC 14


RESULT 2
LANC_STRS6
ID   LANC_STRS6     STANDARD;      PRT;    19 AA.
AC   P38655;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Lantibiotic ancovenin.
OS   Streptomyces sp. (strain A647P-2).
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=72591;
RN   [1]
RP   PROTEIN SEQUENCE.
RA   Wakamiya T., Ueki Y., Shiba T., Kido Y., Motoki Y.;
RT   "The structure of ancovenin, a new peptide inhibitor of angiotensin I
RT   converting enzyme.";
RL   Tetrahedron Lett. 26:665-668(1985).
CC   -!- FUNCTION: Acts as an inhibitor of angiotensin I converting enzyme.
CC   -!- PTM: Maturation of lantibiotics involves the enzymic conversion of
CC       Thr, and Ser into dehydrated AA and the formation of thioether
CC       bonds with cysteine or the formation of dialkylamine bonds with
CC       lysine. This is followed by membrane translocation and cleavage of
CC       the modified precursor.
CC   -!- SIMILARITY: Belongs to the type B lantibiotic family.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   PIR; A61284; EWSMAN.
KW   Antibiotic; Antimicrobial; Bacteriocin; Direct protein sequencing;
KW   Lantibiotic; Thioether bond.
FT   CROSSLNK      1     18       Beta-methyllanthionine (Cys-Thr).
FT   CROSSLNK      4     14       Lanthionine (Ser-Cys).
FT   CROSSLNK      5     11       Beta-methyllanthionine (Cys-Thr).
FT   CROSSLNK      6     19       Lysinoalanine (Ser-Lys).
SQ   SEQUENCE   19 AA;  2033 MW;  F434299E2736286A CRC64;

  Query Match             83.3%;  Score 45;  DB 1;  Length 19;
  Best Local Similarity   60.0%;  Pred. No. 0.26;
  Matches    6;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          4 CXXGPXTWXC 13
              |  ||  || |
Db          5 CSFGPLTWSC 14


RESULT 3
Q8WWV6_HUMAN
ID   Q8WWV6_HUMAN PRELIMINARY;      PRT;   532 AA.
AC   Q8WWV6;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Fc alpha/mu receptor.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=21638011; PubMed=11779189;
RA   McDonald K.J., Cameron A.J.M., Allen J.M., Jardine A.G.;
RT   "Expression of Fc alpha/mu receptor by human mesangial cells: a
RT   candidate receptor for immune complex deposition in IgA nephropathy.";
RL   Biochem. Biophys. Res. Commun. 290:438-442(2002).
DR   EMBL; AY063125; AAL51154.1; -; mRNA.
DR   Ensembl; ENSG00000162897; Homo sapiens.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR007110; Ig-like.
DR   SMART; SM00409; IG; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
KW   Immunoglobulin domain; Receptor.
SQ   SEQUENCE   532 AA;  57144 MW;  D347A23C0F41EED3 CRC64;

  Query Match             75.9%;  Score 41;  DB 2;  Length 532;
  Best Local Similarity   50.0%;  Pred. No. 31;
  Matches    6;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 YXCXXGPXTWXC 13
              | |  ||  | |
Db         96 YWCRLGPPRWIC 107


RESULT 4
Q96SA2_HUMAN
ID   Q96SA2_HUMAN PRELIMINARY;      PRT;   534 AA.
AC   Q96SA2;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   FKSG87 protein.
GN   Name=FKSG87;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Wang Y.-G., Gong L.;
RL   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF354295; AAK39522.1; -; mRNA.
DR   Ensembl; ENSG00000162897; Homo sapiens.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR007110; Ig-like.
DR   SMART; SM00409; IG; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
KW   Immunoglobulin domain.
SQ   SEQUENCE   534 AA;  56749 MW;  6EF8050E412AF91C CRC64;

  Query Match             75.9%;  Score 41;  DB 2;  Length 534;
  Best Local Similarity   50.0%;  Pred. No. 31;
  Matches    6;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 YXCXXGPXTWXC 13
              | |  ||  | |
Db        116 YWCRLGPPRWIC 127


RESULT 5
Q5REH9_PONPY
ID   Q5REH9_PONPY PRELIMINARY;      PRT;   577 AA.
AC   Q5REH9;
DT   01-FEB-2005 (TrEMBLrel. 29, Created)
DT   01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT   01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE   Hypothetical protein DKFZp469K1129.
GN   Name=DKFZp469K1129;
OS   Pongo pygmaeus (Orangutan).
```

```
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Pongo.
OX   NCBI_TaxID=9600;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Kidney;
RG   The German cDNA Consortium;
RA   Ottenwaelder B., Obermaier B., Deutschenbaur S., Schaipp A.,
RA   Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL   Submitted (NOV-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; CR857549; CAH89828.1; -; mRNA.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR007110; Ig-like.
DR   SMART; SM00409; IG; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
KW   Hypothetical protein; Immunoglobulin domain.
SQ   SEQUENCE   577 AA;  62062 MW;  AA0FCEE7AB9C4BCD CRC64;

  Query Match             75.9%;  Score 41;  DB 2;  Length 577;
  Best Local Similarity   50.0%;  Pred. No. 33;
  Matches    6;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 YXCXXGPXTWXC 13
              | |  ||  | |
Db        129 YWCRLGPPRWIC 140


RESULT 6
Q5R770_PONPY
ID   Q5R770_PONPY PRELIMINARY;      PRT;   589 AA.
AC   Q5R770;
DT   01-FEB-2005 (TrEMBLrel. 29, Created)
DT   01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT   01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE   Hypothetical protein DKFZp469A0319.
GN   Name=DKFZp469A0319;
OS   Pongo pygmaeus (Orangutan).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Pongo.
OX   NCBI_TaxID=9600;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Kidney;
RG   The German cDNA Consortium;
RA   Poustka A., Albert R., Moosmayer P., Schupp I., Wellenreuther R.,
RA   Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL   Submitted (NOV-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; CR860248; CAH92390.1; -; mRNA.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR007110; Ig-like.
DR   SMART; SM00409; IG; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
KW   Hypothetical protein; Immunoglobulin domain.
SQ   SEQUENCE   589 AA;  63435 MW;  255BF0FEAACCA812 CRC64;

  Query Match             75.9%;  Score 41;  DB 2;  Length 589;
  Best Local Similarity   50.0%;  Pred. No. 34;
  Matches    6;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 YXCXXGPXTWXC 13
              | |  ||  | |
Db        141 YWCRLGPPRWIC 152


RESULT 7
Q6ZW93_HUMAN
ID   Q6ZW93_HUMAN PRELIMINARY;      PRT;   167 AA.
AC   Q6ZW93;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Hypothetical protein FLJ41423.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Hippocampus;
RA   Kawakami B., Sugiyama A., Takemoto M., Sugiyama T., Irie R.,
RA   Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA   Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA   Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,
RA   Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A., Suzuki Y.,
RA   Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
RL   Submitted (JUL-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK123417; BAC85611.1; -; mRNA.
SQ   SEQUENCE   167 AA;  17960 MW;  266132D59393C276 CRC64;

  Query Match             74.1%;  Score 40;  DB 2;  Length 167;
  Best Local Similarity   50.0%;  Pred. No. 16;
  Matches    5;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          4 CXXGPXTWXC 13
              |  ||  | |
Db         83 CRQGPSVWSC 92


RESULT 8
Q5VHX3_EAV
ID   Q5VHX3_EAV  PRELIMINARY;      PRT;   173 AA.
AC   Q5VHX3;
DT   01-FEB-2005 (TrEMBLrel. 29, Created)
DT   01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT   01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE   Large envelope protein (Fragment).
GN   Name=ORF5;
OS   Equine arteritis virus (EAV).
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC   Arteriviridae; Arterivirus.
OX   NCBI_TaxID=11047;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=S4;
RA   Mittelholzer C., Johansson I., Baule C., Hannant D., Paton D.,
RA   Autorino G.L., Nowotny N., Belak S.;
RT   "Extended phylogeny of equine arteritis virus: division into new
RT   subgroups.";
RL   Submitted (OCT-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY453342; AAS17004.1; -; Genomic_RNA.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   InterPro; IPR001332; Arteri_glycop.
DR   InterPro; IPR003241; EAV_ORF5.
DR   Pfam; PF00951; Arteri_Gl; 1.
DR   ProDom; PD002371; EAV_ORF5; 1.
KW   Envelope protein.
FT   NON_TER       1      1
FT   NON_TER     173    173
SQ   SEQUENCE   173 AA;  19488 MW;  9147CBD1DD750ADE CRC64;

  Query Match             74.1%;  Score 40;  DB 2;  Length 173;
  Best Local Similarity   41.7%;  Pred. No. 17;
  Matches    5;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          2 YXCXXGPXTWXC 13
              | |    |  | ||
Db          5 YNCSASPTCWYC 16


RESULT 9
Q4I355_GIBZE
ID   Q4I355_GIBZE PRELIMINARY;      PRT;   180 AA.
AC   Q4I355;
```

```
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Predicted protein.
GN   ORFNames=FG08353.1;
OS   Gibberella zeae PH-1.
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC   Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
OX   NCBI_TaxID=229533;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=PH-1;
RA   Birren B., Nusbaum C., Abouelleil A., Allen N., Anderson S.,
RA   Arachchi H.M., Barna N., Bastien V., Bloom T., Boguslavkiy L.,
RA   Boukhgalter B., Butler J., Calvo S.E., Camarata J., Chang J.,
RA   Choepel Y., Collymore A., Cook A., Cooke P., Corum B., DeArellano K.,
RA   Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
RA   Erickson J., Faro S., Ferreira P., FitzGerald M., Gage D., Galagan J.,
RA   Gardyna S., Gnerre S., Graham L., Grand-Pierre N., Hafez N.,
RA   Hagopian D., Hagos B., Hall J., Horton L., Hulme W., Iliev I.,
RA   Jaffe D., Johnson R., Jones C., Kamal M., Kamat A., Karatas A.,
RA   Kells C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
RA   Ma L.-J., Mabbitt R., MacLean C., Macdonald P., Major J., Manning J.,
RA   Matthews C., Mauceli E., McCarthy M., Meldrim J., Meneus L.,
RA   Mihova T., Mlenga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
RA   Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neil D.,
RA   Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,
RA   Rachupka A., Ramasamy U., Raymond C., Retta R., Rise C., Rogov P.,
RA   Roman J., Schauer S., Schupback R., Seaman S., Severy P., Smirnov S.,
RA   Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,
RA   Talamas J., Tesfaye S., Theodore J., Topham K., Travers M.,
RA   Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,
RA   Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
RA   Lander E.;
RT   "Fusarium graminearum genome sequence.";
RL   Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; AACM01000335; EAA72141.1; -; Genomic_DNA.
SQ   SEQUENCE   180 AA;  20463 MW;  94C7B5242FEE6ED9 CRC64;

  Query Match             74.1%;  Score 40;  DB 2;  Length 180;
  Best Local Similarity   41.7%;  Pred. No. 17;
  Matches    5;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 YXCXXGPXTWXC 13
              : |  ||  | |
Db         99 HNCSFGPAPWEC 110


RESULT 10
Q4IMN3_GIBZE
ID   Q4IMN3_GIBZE PRELIMINARY;      PRT;   352 AA.
AC   Q4IMN3;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Hypothetical protein.
GN   ORFNames=FG01525.1;
OS   Gibberella zeae PH-1.
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC   Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
OX   NCBI_TaxID=229533;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=PH-1;
RA   Birren B., Nusbaum C., Abouelleil A., Allen N., Anderson S.,
RA   Arachchi H.M., Barna N., Bastien V., Bloom T., Boguslavkiy L.,
RA   Boukhgalter B., Butler J., Calvo S.E., Camarata J., Chang J.,
RA   Choepel Y., Collymore A., Cook A., Cooke P., Corum B., DeArellano K.,
RA   Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
RA   Erickson J., Faro S., Ferreira P., FitzGerald M., Gage D., Galagan J.,
RA   Gardyna S., Gnerre S., Graham L., Grand-Pierre N., Hafez N.,
RA   Hagopian D., Hagos B., Hall J., Horton L., Hulme W., Iliev I.,
RA   Jaffe D., Johnson R., Jones C., Kamal M., Kamat A., Karatas A.,
RA   Kells C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
RA   Ma L.-J., Mabbitt R., MacLean C., Macdonald P., Major J., Manning J.,
RA   Matthews C., Mauceli E., McCarthy M., Meldrim J., Meneus L.,
RA   Mihova T., Mlenga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
RA   Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neil D.,
RA   Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,
RA   Rachupka A., Ramasamy U., Raymond C., Retta R., Rise C., Rogov P.,
RA   Roman J., Schauer S., Schupback R., Seaman S., Severy P., Smirnov S.,
RA   Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,
RA   Talamas J., Tesfaye S., Theodore J., Topham K., Travers M.,
RA   Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,
RA   Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
RA   Lander E.;
RT   "Fusarium graminearum genome sequence.";
RL   Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; AACM01000077; EAA68151.1; -; Genomic_DNA.
KW   Hypothetical protein.
SQ   SEQUENCE   352 AA;  38308 MW;  670BA49FC645A7F8 CRC64;

  Query Match             74.1%;  Score 40;  DB 2;  Length 352;
  Best Local Similarity   50.0%;  Pred. No. 32;
  Matches    5;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          4 CXXGPXTWXC 13
              |   | || |
Db        184 CTSNPSTWRC 193


RESULT 11
Q61J27_CAEBR
ID   Q61J27_CAEBR PRELIMINARY;      PRT;  2022 AA.
AC   Q61J27;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Hypothetical protein CBG09974 (Fragment).
GN   Name=CBG09974;
OS   Caenorhabditis briggsae.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6238;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RG   The C.briggsae Sequencing Consortium;
RL   Submitted (SEP-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; CAAC01000045; CAE65110.1; -; Genomic_DNA.
DR   SMR; Q61J27; 403-514, 600-683, 1553-1635.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0005515; F:protein binding; IEA.
DR   GO; GO:0003713; F:transcription coactivator activity; IEA.
DR   GO; GO:0008270; F:zinc ion binding; IEA.
DR   GO; GO:0007049; P:cell cycle; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR001487; Bromodomain.
DR   InterPro; IPR010303; DUF902_CREBbp.
DR   InterPro; IPR003101; KIX.
DR   InterPro; IPR000197; TAZ_finger.
DR   InterPro; IPR009255; Trans_coact.
DR   InterPro; IPR001965; Znf_PHD.
DR   InterPro; IPR000433; Znf_ZZ.
DR   Pfam; PF00439; Bromodomain; 1.
DR   Pfam; PF06001; DUF902; 1.
DR   Pfam; PF06010; DUF906; 1.
DR   Pfam; PF02172; KIX; 1.
DR   Pfam; PF02135; zf-TAZ; 2.
DR   Pfam; PF00569; ZZ; 1.
DR   PRINTS; PR00503; BROMODOMAIN.
```

```
DR   SMART; SM00297; BROMO; 1.
DR   SMART; SM00551; ZnF_TAZ; 2.
DR   SMART; SM00291; ZnF_ZZ; 1.
DR   PROSITE; PS00633; BROMODOMAIN_1; 1.
DR   PROSITE; PS50014; BROMODOMAIN_2; 1.
DR   PROSITE; PS50952; KIX; 1.
DR   PROSITE; PS01359; ZF_PHD_1; UNKNOWN_1.
DR   PROSITE; PS50134; ZF_TAZ; 2.
DR   PROSITE; PS50135; ZF_ZZ_2; 1.
KW   Hypothetical protein.
FT   NON_TER    2022   2022
SQ   SEQUENCE   2022 AA;  222664 MW;  6C7C9A621C30D950 CRC64;

  Query Match             74.1%;  Score 40;  DB 2;  Length 2022;
  Best Local Similarity   50.0%;  Pred. No. 1.6e+02;
  Matches    7;  Conservative    0;  Mismatches    5;  Indels    2;  Gaps    1;

Qy          2 YXCX--XGPXTWXC 13
              | |    || || |
Db       1499 YTCNKCNGPATWHC 1512


RESULT 12
Q6ZT75_HUMAN
ID   Q6ZT75_HUMAN PRELIMINARY;      PRT;   172 AA.
AC   Q6ZT75;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Hypothetical protein FLJ44897.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Amygdala;
RA   Oshima A., Takahashi-Fujii A., Tanase T., Imose N., Takeuchi K.,
RA   Arita M., Musashino K., Yuuki H., Hara H., Sugiyama T., Irie R.,
RA   Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA   Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA   Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,
RA   Murakawa K., Kanehori K., Sugiyama A., Kawakami B., Suzuki Y.,
RA   Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
RL   Submitted (JUL-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK126845; BAC86719.1; -; mRNA.
SQ   SEQUENCE   172 AA;  18807 MW;  DFD5579875B25559 CRC64;

  Query Match             72.2%;  Score 39;  DB 2;  Length 172;
  Best Local Similarity   50.0%;  Pred. No. 25;
  Matches    5;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          4 CXXGPXTWXC 13
              |  || :| |
Db          9 CLCGPESWTC 18


RESULT 13
Q6ZWC2_HUMAN
ID   Q6ZWC2_HUMAN PRELIMINARY;      PRT;   172 AA.
AC   Q6ZWC2;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Hypothetical protein FLJ41341.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Brain;
RA   Tanigami A., Fujiwara T., Shibahara T., Goto Y., Hirao M., Shimizu F.,
RA   Wakebe H., Ono T., Hishigaki H., Watanabe T., Ozaki K., Sugiyama T.,
RA   Irie R., Otsuki T., Sato H., Ota T., Wakamatsu A., Ishii S.,
RA   Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
RA   Kimura K., Yamashita H., Matsuo K., Nakamura Y., Sekine M.,
RA   Kikuchi H., Kanda K., Wagatsuma M., Murakawa K., Kanehori K.,
RA   Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B., Suzuki Y.,
RA   Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
RL   Submitted (JUL-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK123335; BAC85582.1; -; mRNA.
SQ   SEQUENCE   172 AA;  18777 MW;  C565579875A8FFF8 CRC64;

  Query Match             72.2%;  Score 39;  DB 2;  Length 172;
  Best Local Similarity   50.0%;  Pred. No. 25;
  Matches    5;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          4 CXXGPXTWXC 13
              |  || :| |
Db          9 CLCGPESWTC 18


RESULT 14
Q9WD22_EAV
ID   Q9WD22_EAV  PRELIMINARY;      PRT;   173 AA.
AC   Q9WD22;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Large envelope protein (Fragment).
GN   Name=ORF5;
OS   Equine arteritis virus (EAV).
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC   Arteriviridae; Arterivirus.
OX   NCBI_TaxID=11047;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=470VE1/95;
RX   MEDLINE=99190430; PubMed=10092009;
RA   Stadejek T., Bjorklund H., Bascunana C.R., Ciabatti I.M.,
RA   Scicluna M.T., Amaddeo D., McCollum W.H., Autorino G.L., Timoney P.J.,
RA   Paton D.J., Klingeborn B., Belak S.;
RT   "Genetic diversity of equine arteritis virus.";
RL   J. Gen. Virol. 80:691-699(1999).
DR   EMBL; AF099825; AAD24933.1; -; Genomic_RNA.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   InterPro; IPR001332; Arteri_glycop.
DR   InterPro; IPR003241; EAV_ORF5.
DR   Pfam; PF00951; Arteri_Gl; 1.
DR   ProDom; PD002371; EAV_ORF5; 1.
KW   Envelope protein.
FT   NON_TER       1      1
FT   NON_TER     173    173
SQ   SEQUENCE   173 AA;  19386 MW;  D701BE129DC62E59 CRC64;

  Query Match             72.2%;  Score 39;  DB 2;  Length 173;
  Best Local Similarity   41.7%;  Pred. No. 25;
  Matches    5;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          2 YXCXXGPXTWXC 13
              | |   |  | ||
Db          5 YNCSASPNCWYC 16


RESULT 15
O70227_RAT
ID   O70227_RAT  PRELIMINARY;      PRT;    61 AA.
AC   O70227;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-AUG-1998 (TrEMBLrel. 07, Last annotation update)
DE   MARRLC9A (Fragment).
OS   Rattus norvegicus (Rat).
```

```
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC   Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Brain;
RA   Liao B.S., Jin W.L., Ju G.;
RL   Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF010444; AAC14892.1; -; mRNA.
FT   NON_TER       1      1
FT   NON_TER      61     61
SQ   SEQUENCE   61 AA;  6655 MW;  C8AF3B9CB8656126 CRC64;

  Query Match             70.4%;  Score 38;  DB 2;  Length 61;
  Best Local Similarity   41.7%;  Pred. No. 15;
  Matches    5;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 YXCXXGPXTWXC 13
              : |  ||  | |
Db         21 HLCPRGPGQWAC 32


Search completed: March 31, 2006, 16:35:16
Job time : 59.9403 secs
```

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: March 31, 2006, 16:09:06 ; Search time 38.5572 Seconds
(without alignments)
113.955 Million cell updates/sec

Title: US-10-609-217-421
Perfect score: 47
Sequence: 1 CXXGPXTWXC 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	45	95.7	16	2	AAR62636	Aar62636 Intermedi
2	45	95.7	19	8	ADN11733	Adn11733 Streptomy
3	45	95.7	19	8	ADN11732	Adn11732 Streptomy
4	45	95.7	21	9	ADU91978	Adu91978 EPO-R ago
5	45	95.7	24	2	AAY26477	Aay26477 Erythropo
6	45	95.7	24	2	AAY26424	Aay26424 Erythropo
7	45	95.7	24	2	AAY26431	Aay26431 Erythropo
8	44	93.6	17	9	ADU91962	Adu91962 EPO-R ago
9	44	93.6	20	2	AAY13704	Aay13704 Erythropo
10	44	93.6	20	2	AAY26527	Aay26527 Erythropo
11	44	93.6	20	2	AAY13696	Aay13696 Erythropo
12	44	93.6	20	2	AAY13650	Aay13650 Erythropo
13	44	93.6	20	2	AAY13728	Aay13728 Erythropo
14	44	93.6	20	2	AAY13688	Aay13688 Erythropo
15	44	93.6	20	2	AAY13687	Aay13687 Erythropo
16	44	93.6	20	2	AAY13705	Aay13705 Erythropo
17	44	93.6	20	2	AAY26368	Aay26368 Erythropo
18	44	93.6	20	2	AAY13672	Aay13672 Erythropo
19	44	93.6	20	2	AAY13689	Aay13689 Erythropo
20	44	93.6	20	2	AAY13706	Aay13706 Erythropo
21	44	93.6	20	2	AAY13679	Aay13679 Erythropo
22	44	93.6	20	2	AAY13727	Aay13727 Erythropo
23	44	93.6	20	2	AAY13662	Aay13662 Erythropo
24	44	93.6	20	2	AAW27019	Aaw27019 Monomer s
25	44	93.6	20	2	AAW27001	Aaw27001 Monomer s
26	44	93.6	20	2	AAW27041	Aaw27041 Monomer s
27	44	93.6	20	2	AAW26993	Aaw26993 Monomer s
28	44	93.6	20	2	AAW27020	Aaw27020 Monomer s
29	44	93.6	20	2	AAW27010	Aaw27010 Monomer s
30	44	93.6	20	2	AAW27003	Aaw27003 Monomer s
31	44	93.6	20	2	AAW26976	Aaw26976 Monomer s
32	44	93.6	20	2	AAW27018	Aaw27018 Monomer s
33	44	93.6	20	2	AAW27002	Aaw27002 Monomer s
34	44	93.6	20	2	AAW26986	Aaw26986 Monomer s
35	44	93.6	20	5	AAU74480	Aau74480 Human ery
36	44	93.6	22	2	AAY13709	Aay13709 Erythropo
37	44	93.6	22	2	AAY26525	Aay26525 Erythropo
38	44	93.6	22	2	AAY26491	Aay26491 Erythropo
39	44	93.6	22	2	AAY26511	Aay26511 Erythropo
40	44	93.6	22	2	AAY26415	Aay26415 Erythropo
41	44	93.6	22	2	AAY26355	Aay26355 Erythropo
42	44	93.6	22	2	AAY26382	Aay26382 Erythropo
43	44	93.6	22	2	AAY26528	Aay26528 Erythropo
44	44	93.6	22	2	AAY26502	Aay26502 Erythropo
45	44	93.6	22	2	AAW27023	Aaw27023 Monomer s

ALIGNMENTS

RESULT 1
AAR62636
ID AAR62636 standard; peptide; 16 AA.
XX
AC AAR62636;
XX
DT 15-JUN-1995 (first entry)
XX
DE Intermediate for synthesis of lanthionine and methyllanthionine.
XX
KW Lanthionine; methyllanthionine; lantibiotic; antiviral;
KW immunosuppressant; antimicrobial; enzyme inhibitor.
XX
OS Synthetic.
XX
PN JP06253885-A.
XX
PD 13-SEP-1994.
XX
PF 09-MAR-1993; 93JP-00048385.
XX
PR 09-MAR-1993; 93JP-00048385.
XX
PA (AJIN) AJINOMOTO KK.
XX
DR WPI; 1994-329026/41.
XX
PT Prepn. of lantionine contg. peptide(s) - useful as antimicrobial,
PT antiviral drugs, immunosuppressants and enzyme inhibitors.
XX
PS Example 2; Page 7; 8pp; Japanese.
XX
CC This is one of 5 peptides (AAR62635-R63639) containing Cys and Ser or Thr
CC residues which were synthesised and tested for their usefulness as
CC intermediates for the preparation of peptides which include lanthionine.
CC Peptides 3 and 4 (AAR62637-8) produced lanthionine, while both
CC lanthionine and methyllanthionine could be produced from peptide 2
CC (AAR62636)
XX
SQ Sequence 16 AA;

Query Match 95.7%; Score 45; DB 2; Length 16;
Best Local Similarity 60.0%; Pred. No. 1.2;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CXXGPXTWXC 10
| || || |

```
Db          5 CSFGPLTWSC 14


RESULT 2
ADN11733
ID   ADN11733 standard; protein; 19 AA.
XX
AC   ADN11733;
XX
DT   15-JUL-2004  (first entry)
XX
DE   Streptomyces ancovenin propeptide.
XX
KW   duramycin; antibiotic; gene therapy; preduramycin; produramycin.
XX
OS   Streptomyces sp.
XX
PN   WO2004033706-A2.
XX
PD   22-APR-2004.
XX
PF   22-SEP-2003; 2003WO-US029852.
XX
PR   10-OCT-2002; 2002US-0417709P.
XX
PA   (MOLI-) MOLICHEM MEDICINES INC.
XX
PI   Molina L,  Romeo CJ;
XX
DR   WPI; 2004-340939/31.
XX
PT   New nucleic acid comprising a sequence encoding preduramycin or
PT   produramycin, useful in making preduramycin, produramycin or duramycin.
XX
PS   Example 1; Page 26; 28pp; English.
XX
CC   The present invention provides the coding sequence of the Streptomyces
CC   cinnamoneus antibiotic preduramycin. The nucleic acid is useful in making
CC   preduramycin, produramycin or duramycin. The present sequence is a
CC   polypeptide fragment of the invention.
XX
SQ   Sequence 19 AA;

  Query Match             95.7%;  Score 45;  DB 8;  Length 19;
  Best Local Similarity   60.0%;  Pred. No. 1.3;
  Matches    6;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 CXXGPXTWXC 10
              |  || || |
Db          5 CSFGPLTWSC 14


RESULT 3
ADN11732
ID   ADN11732 standard; protein; 19 AA.
XX
AC   ADN11732;
XX
DT   15-JUL-2004  (first entry)
XX
DE   Streptomyces duramycin C propeptide.
XX
KW   duramycin; antibiotic; gene therapy; preduramycin; produramycin.
XX
OS   Streptomyces sp.
XX
PN   WO2004033706-A2.
XX
PD   22-APR-2004.
XX
PF   22-SEP-2003; 2003WO-US029852.
XX
PR   10-OCT-2002; 2002US-0417709P.
XX
PA   (MOLI-) MOLICHEM MEDICINES INC.
XX
PI   Molina L,  Romeo CJ;
XX
DR   WPI; 2004-340939/31.
XX
PT   New nucleic acid comprising a sequence encoding preduramycin or
PT   produramycin, useful in making preduramycin, produramycin or duramycin.
XX
PS   Example 1; Page 25-26; 28pp; English.
XX
CC   The present invention provides the coding sequence of the Streptomyces
CC   cinnamoneus antibiotic preduramycin. The nucleic acid is useful in making
CC   preduramycin, produramycin or duramycin. The present sequence is a
CC   polypeptide fragment of the invention.
XX
SQ   Sequence 19 AA;

  Query Match             95.7%;  Score 45;  DB 8;  Length 19;
  Best Local Similarity   60.0%;  Pred. No. 1.3;
  Matches    6;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 CXXGPXTWXC 10
              |  || || |
Db          5 CSYGPLTWSC 14


RESULT 4
ADU91978
ID   ADU91978 standard; peptide; 21 AA.
XX
AC   ADU91978;
XX
DT   10-FEB-2005  (first entry)
XX
DE   EPO-R agonist SEQ ID NO 119.
XX
KW   erythropoietin receptor; EPO-R; erythropoietin; renal failure;
KW   autoimmune disease; cystic fibrosis; anemia; inflammation;
KW   spinal cord injury; aging; neurological disease; nephrotropic;
KW   antianemic; immunosuppressive; CNS-Gen.; neuroprotective;
KW   respiratory-Gen.; antiinflammatory; vulnerary; nootropic; cytostatic;
KW   hemostatic; cyclic.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "Acetylated residue"
FT   Disulfide-bond  7. .16
FT   Modified-site   21
FT                   /note= "C-terminal amide"
XX
PN   WO2004101611-A2.
XX
PD   25-NOV-2004.
XX
PF   12-MAY-2004; 2004WO-US014886.
XX
PR   12-MAY-2003; 2003US-0470245P.
XX
PA   (AFFY-) AFFYMAX INC.
XX
PI   Yin K,  Holmes C,  Lalonde G,  Balu P,  Schatz PJ,  Tumelty D;
XX
DR   WPI; 2005-039329/04.
XX
PT   New peptide comprising specified sequence of amino acid is erythropoietin
PT   receptor agonist useful for treating e.g. anemia, beta-thalassemia, renal
PT   disorders.
XX
PS   Disclosure; SEQ ID NO 119; 83pp; English.
```

```
XX
CC   This invention describes a novel peptide which is an erythropoietin
CC   receptor (EPO-R) activator. The peptide forms a dimer comprising a
CC   linking moiety connecting two peptide chains composed of ADU91861. The N-
CC   terminal of the peptide is acetylated. The EPO-R activator further
CC   comprises at least one water soluble polymer, preferably polyethylene
CC   glycol (PEG) covalently bound to the peptide and a spacer moiety. The
CC   products of the invention are used for treating disorders associated with
CC   deficiency of erythropoietin or low or defective red blood cell
CC   population, end stage renal failure or dialysis, anemia associated with
CC   AIDS, autoimmune disease or malignancy, beta-thalassemia, cystic
CC   fibrosis, early anemia of prematurity, anemia associated with chronic
CC   inflammatory disease, spinal cord injury, acute blood loss, aging and
CC   neoplastic disease states accompanied by abnormal erythropoiesis. The
CC   peptide compounds are potent agonists of erythropoietin receptor and have
CC   nephrotropic, antianemic, immunosuppressive, CNS-Gen., neuroprotective,
CC   respiratory-Gen., antiinflammatory, vulnerary, nootropic, cytostatic and
CC   hemostatic activity. This sequence represents a peptide which acts as an
CC   erythropoietin receptor (EPO-R) agonist.
XX
SQ   Sequence 21 AA;

  Query Match             95.7%;  Score 45;  DB 9;  Length 21;
  Best Local Similarity   60.0%;  Pred. No. 1.5;
  Matches    6;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 CXXGPXTWXC 10
              |  ||  || |
Db          7 CQMGPTTWTC 16


RESULT 5
AAY26477
ID   AAY26477 standard; peptide; 24 AA.
XX
AC   AAY26477;
XX
DT   06-SEP-1999  (first entry)
XX
DE   Erythropoietin receptor (EPO-R) binding peptide.
XX
KW   Erythropoietin; EPO; receptor; EPO deficiency; renal failure; AIDS;
KW   dialysis; anaemia; autoimmune disease; chronic inflammatory disease;
KW   malignancy; beta-thalassaemia; cystic fibrosis; prematurity; blood loss;
KW   spinal cord injury; aging; neoplastic disease; erythropoiesis; EPO-R.
XX
OS   Synthetic.
XX
PN   WO9640749-A1.
XX
PD   19-DEC-1996.
XX
PF   07-JUN-1996;   96WO-US009810.
XX
PR   07-JUN-1995;   95US-00484631.
PR   07-JUN-1995;   95US-00484635.
XX
PA   (JOHJ ) JOHNSON & JOHNSON CORP.
PA   (AFFY-) AFFYMAX TECHNOLOGIES NV.
XX
PI   Wrighton NC,  Dower WJ,  Chang RS,  Kashyap AK,  Jolliffe LK;
PI   Johnson D,  Mulcahy L;
XX
DR   WPI; 1997-052225/05.
XX
PT   Erythropoietin receptor binding peptide - useful for treating disorders
PT   characterised by deficiency of EPO, or low or defective red blood cell
PT   population.
XX
PS   Disclosure; Page 22; 95pp; English.
XX
CC   The invention describes a peptide of 10-40 amino acid residues which
CC   binds to erythropoietin (EPO) receptor and which includes the amino acid
CC   sequence Cys-Xaa1-Xaa2-Gly-Pro-Xaa3-Thr-Trp-Xaa4-Cys, where Xaa1 = Arg,
CC   His, Leu or Trp, Xaa2 = Met, Phe or Ile, Xaa3 = 1 of the 20 genetically
CC   coded L-amino acids, and Xaa4 = Asp, Glu, Ile, Leu or Val. Optionally,
CC   the peptide may be cyclised or dimerised. The peptide can be used to
CC   treat a patient having a disorder characterised by a deficiency of EPO or
CC   a low or defective red blood cell population. It can be used to treat end
CC   stage renal failure or dialysis; anaemia associated with AIDS, autoimmune
CC   disease, chronic inflammatory diseases or malignancy; beta-thalassaemia;
CC   cystic fibrosis; early anaemia of prematurity; spinal cord injury; acute
CC   blood loss; aging; and neoplastic disease states accompanied by abnormal
CC   erythropoiesis. The peptides can also be used as reagents for detecting
CC   EPO receptors on living cells, in biological fluids, in tissue
CC   homogenates, etc. Sequences AAY26352-548 are representative peptides
CC   falling within the above peptide motif and isolated during the affinity
CC   selection process
XX
SQ   Sequence 24 AA;

  Query Match             95.7%;  Score 45;  DB 2;  Length 24;
  Best Local Similarity   60.0%;  Pred. No. 1.6;
  Matches    6;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 CXXGPXTWXC 10
              |  ||  || |
Db          8 CSKGPATWKC 17


RESULT 6
AAY26424
ID   AAY26424 standard; peptide; 24 AA.
XX
AC   AAY26424;
XX
DT   06-SEP-1999  (first entry)
XX
DE   Erythropoietin receptor (EPO-R) binding peptide.
XX
KW   Erythropoietin; EPO; receptor; EPO deficiency; renal failure; AIDS;
KW   dialysis; anaemia; autoimmune disease; chronic inflammatory disease;
KW   malignancy; beta-thalassaemia; cystic fibrosis; prematurity; blood loss;
KW   spinal cord injury; aging; neoplastic disease; erythropoiesis; EPO-R.
XX
OS   Synthetic.
XX
PN   WO9640749-A1.
XX
PD   19-DEC-1996.
XX
PF   07-JUN-1996;   96WO-US009810.
XX
PR   07-JUN-1995;   95US-00484631.
PR   07-JUN-1995;   95US-00484635.
XX
PA   (JOHJ ) JOHNSON & JOHNSON CORP.
PA   (AFFY-) AFFYMAX TECHNOLOGIES NV.
XX
PI   Wrighton NC,  Dower WJ,  Chang RS,  Kashyap AK,  Jolliffe LK;
PI   Johnson D,  Mulcahy L;
XX
DR   WPI; 1997-052225/05.
XX
PT   Erythropoietin receptor binding peptide - useful for treating disorders
PT   characterised by deficiency of EPO, or low or defective red blood cell
PT   population.
XX
PS   Disclosure; Page 20; 95pp; English.
XX
CC   The invention describes a peptide of 10-40 amino acid residues which
CC   binds to erythropoietin (EPO) receptor and which includes the amino acid
CC   sequence Cys-Xaa1-Xaa2-Gly-Pro-Xaa3-Thr-Trp-Xaa4-Cys, where Xaa1 = Arg,
CC   His, Leu or Trp, Xaa2 = Met, Phe or Ile, Xaa3 = 1 of the 20 genetically
CC   coded L-amino acids, and Xaa4 = Asp, Glu, Ile, Leu or Val. Optionally,
CC   the peptide may be cyclised or dimerised. The peptide can be used to
```

```
CC   treat a patient having a disorder characterised by a deficiency of EPO or
CC   a low or defective red blood cell population. It can be used to treat end
CC   stage renal failure or dialysis; anaemia associated with AIDS, autoimmune
CC   disease, chronic inflammatory diseases or malignancy; beta-thalassaemia;
CC   cystic fibrosis; early anaemia of prematurity; spinal cord injury; acute
CC   blood loss; aging; and neoplastic disease states accompanied by abnormal
CC   erythropoiesis. The peptides can also be used as reagents for detecting
CC   EPO receptors on living cells, in biological fluids, in tissue
CC   homogenates, etc. Sequences AAY26352-548 are representative peptides
CC   falling within the above peptide motif and isolated during the affinity
CC   selection process
XX
SQ   Sequence 24 AA;

  Query Match             95.7%;  Score 45;  DB 2;  Length 24;
  Best Local Similarity   60.0%;  Pred. No. 1.6;
  Matches    6;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 CXXGPXTWXC 10
              |  ||  || |
Db         13 CSEGPATWLC 22


RESULT 7
AAY26431
ID   AAY26431 standard; peptide; 24 AA.
XX
AC   AAY26431;
XX
DT   06-SEP-1999  (first entry)
XX
DE   Erythropoietin receptor (EPO-R) binding peptide.
XX
KW   Erythropoietin; EPO; receptor; EPO deficiency; renal failure; AIDS;
KW   dialysis; anaemia; autoimmune disease; chronic inflammatory disease;
KW   malignancy; beta-thalassaemia; cystic fibrosis; prematurity; blood loss;
KW   spinal cord injury; aging; neoplastic disease; erythropoiesis; EPO-R.
XX
OS   Synthetic.
XX
PN   WO9640749-A1.
XX
PD   19-DEC-1996.
XX
PF   07-JUN-1996;   96WO-US009810.
XX
PR   07-JUN-1995;   95US-00484631.
PR   07-JUN-1995;   95US-00484635.
XX
PA   (JOHJ ) JOHNSON & JOHNSON CORP.
PA   (AFFY-) AFFYMAX TECHNOLOGIES NV.
XX
PI   Wrighton NC,  Dower WJ,  Chang RS,  Kashyap AK,  Jolliffe LK;
PI   Johnson D,  Mulcahy L;
XX
DR   WPI; 1997-052225/05.
XX
PT   Erythropoietin receptor binding peptide - useful for treating disorders
PT   characterised by deficiency of EPO, or low or defective red blood cell
PT   population.
XX
PS   Disclosure; Page 20; 95pp; English.
XX
CC   The invention describes a peptide of 10-40 amino acid residues which
CC   binds to erythropoietin (EPO) receptor and which includes the amino acid
CC   sequence Cys-Xaa1-Xaa2-Gly-Pro-Xaa3-Thr-Trp-Xaa4-Cys, where Xaa1 = Arg,
CC   His, Leu or Trp, Xaa2 = Met, Phe or Ile, Xaa3 = 1 of the 20 genetically
CC   coded L-amino acids, and Xaa4 = Asp, Glu, Ile, Leu or Val. Optionally,
CC   the peptide may be cyclised or dimerised. The peptide can be used to
CC   treat a patient having a disorder characterised by a deficiency of EPO or
CC   a low or defective red blood cell population. It can be used to treat end
CC   stage renal failure or dialysis; anaemia associated with AIDS, autoimmune
CC   disease, chronic inflammatory diseases or malignancy; beta-thalassaemia;
CC   cystic fibrosis; early anaemia of prematurity; spinal cord injury; acute
CC   blood loss; aging; and neoplastic disease states accompanied by abnormal
CC   erythropoiesis. The peptides can also be used as reagents for detecting
CC   EPO receptors on living cells, in biological fluids, in tissue
CC   homogenates, etc. Sequences AAY26352-548 are representative peptides
CC   falling within the above peptide motif and isolated during the affinity
CC   selection process
XX
SQ   Sequence 24 AA;

  Query Match             95.7%;  Score 45;  DB 2;  Length 24;
  Best Local Similarity   60.0%;  Pred. No. 1.6;
  Matches    6;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 CXXGPXTWXC 10
              |  ||  || |
Db         13 CSIGPSTWRC 22


RESULT 8
ADU91962
ID   ADU91962 standard; peptide; 17 AA.
XX
AC   ADU91962;
XX
DT   10-FEB-2005  (first entry)
XX
DE   EPO-R agonist SEQ ID NO 103.
XX
KW   erythropoietin receptor; EPO-R; erythropoietin; renal failure;
KW   autoimmune disease; cystic fibrosis; anemia; inflammation;
KW   spinal cord injury; aging; neurological disease; nephrotropic;
KW   antianemic; immunosuppressive; CNS-Gen.; neuroprotective;
KW   respiratory-Gen.; antiinflammatory; vulnerary; nootropic; cytostatic;
KW   hemostatic; cyclic.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "Acetylated residue"
FT   Disulfide-bond  4. .13
FT   Modified-site   17
FT                   /note= "C-terminal amide"
XX
PN   WO2004101611-A2.
XX
PD   25-NOV-2004.
XX
PF   12-MAY-2004; 2004WO-US014886.
XX
PR   12-MAY-2003; 2003US-0470245P.
XX
PA   (AFFY-) AFFYMAX INC.
XX
PI   Yin K,  Holmes C,  Lalonde G,  Balu P,  Schatz PJ,  Tumelty D;
XX
DR   WPI; 2005-039329/04.
XX
PT   New peptide comprising specified sequence of amino acid is erythropoietin
PT   receptor agonist useful for treating e.g. anemia, beta-thalassemia, renal
PT   disorders.
XX
PS   Disclosure; SEQ ID NO 103; 83pp; English.
XX
CC   This invention describes a novel peptide which is an erythropoietin
CC   receptor (EPO-R) activator. The peptide forms a dimer comprising a
CC   linking moiety connecting two peptide chains composed of ADU91861. The N-
CC   terminal of the peptide is acetylated. The EPO-R activator further
CC   comprises at least one water soluble polymer, preferably polyethylene
CC   glycol (PEG) covalently bound to the peptide and a spacer moiety. The
CC   products of the invention are used for treating disorders associated with
CC   deficiency of erythropoietin or low or defective red blood cell
```

```
CC   population, end stage renal failure or dialysis, anemia associated with
CC   AIDS, autoimmune disease or malignancy, beta-thalassemia, cystic
CC   fibrosis, early anemia of prematurity, anemia associated with chronic
CC   inflammatory disease, spinal cord injury, acute blood loss, aging and
CC   neoplastic disease states accompanied by abnormal erythropoiesis. The
CC   peptide compounds are potent agonists of erythropoietin receptor and have
CC   nephrotropic, antianemic, immunosuppressive, CNS-Gen., neuroprotective,
CC   respiratory-Gen., antiinflammatory, vulnerary, nootropic, cytostatic and
CC   hemostatic activity. This sequence represents a peptide which acts as an
CC   erythropoietin receptor (EPO-R) agonist.
XX
SQ   Sequence 17 AA;

  Query Match             93.6%;  Score 44;  DB 9;  Length 17;
  Best Local Similarity   60.0%;  Pred. No. 1.8;
  Matches    6;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 CXXGPXTWXC 10
              |  ||  ||  |
Db          4 CRRGPTTWLC 13


RESULT 9
AAY13704
ID   AAY13704 standard; peptide; 20 AA.
XX
AC   AAY13704;
XX
DT   06-SEP-1999  (first entry)
XX
DE   Erythropoietin receptor (EPO-R) binding peptide.
XX
KW   Erythropoietin; EPO; receptor; EPO deficiency; renal failure; AIDS;
KW   dialysis; anaemia; autoimmune disease; chronic inflammatory disease;
KW   malignancy; beta-thalassaemia; cystic fibrosis; prematurity; blood loss;
KW   spinal cord injury; aging; neoplastic disease; erythropoiesis; EPO-R.
XX
OS   Synthetic.
XX
PN   WO9640749-A1.
XX
PD   19-DEC-1996.
XX
PF   07-JUN-1996;   96WO-US009810.
XX
PR   07-JUN-1995;   95US-00484631.
PR   07-JUN-1995;   95US-00484635.
XX
PA   (JOHJ ) JOHNSON & JOHNSON CORP.
PA   (AFFY-) AFFYMAX TECHNOLOGIES NV.
XX
PI   Wrighton NC,  Dower WJ,  Chang RS,  Kashyap AK,  Jolliffe LK;
PI   Johnson D,  Mulcahy L;
XX
DR   WPI; 1997-052225/05.
XX
PT   Erythropoietin receptor binding peptide - useful for treating disorders
PT   characterised by deficiency of EPO, or low or defective red blood cell
PT   population.
XX
PS   Disclosure; Fig 2; 95pp; English.
XX
CC   The invention describes a peptide of 10-40 amino acid residues which
CC   binds to erythropoietin (EPO) receptor and which includes the amino acid
CC   sequence Cys-Xaa1-Xaa2-Gly-Pro-Xaa3-Thr-Trp-Xaa4-Cys, where Xaa1 = Arg,
CC   His, Leu or Trp, Xaa2 = Met, Phe or Ile, Xaa3 = 1 of the 20 genetically
CC   coded L-amino acids, and Xaa4 = Asp, Glu, Ile, Leu or Val. Optionally,
CC   the peptide may be cyclised or dimerised. The peptide can be used to
CC   treat a patient having a disorder characterised by a deficiency of EPO or
CC   a low or defective red blood cell population. It can be used to treat end
CC   stage renal failure or dialysis; anaemia associated with AIDS, autoimmune
CC   disease, chronic inflammatory diseases or malignancy; beta-thalassaemia;
CC   cystic fibrosis; early anaemia of prematurity; spinal cord injury; acute
CC   blood loss; aging; and neoplastic disease states accompanied by abnormal
CC   erythropoiesis. The peptides can also be used as reagents for detecting
CC   EPO receptors on living cells, in biological fluids, in tissue
CC   homogenates, etc. Sequences AAY13662-735 are representative peptides of
CC   the invention
XX
SQ   Sequence 20 AA;

  Query Match             93.6%;  Score 44;  DB 2;  Length 20;
  Best Local Similarity   60.0%;  Pred. No. 2.1;
  Matches    6;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 CXXGPXTWXC 10
              |  ||  ||  |
Db          6 CRFGPTTWLC 15


RESULT 10
AAY26527
ID   AAY26527 standard; peptide; 20 AA.
XX
AC   AAY26527;
XX
DT   06-SEP-1999  (first entry)
XX
DE   Erythropoietin receptor (EPO-R) binding peptide.
XX
KW   Erythropoietin; EPO; receptor; EPO deficiency; renal failure; AIDS;
KW   dialysis; anaemia; autoimmune disease; chronic inflammatory disease;
KW   malignancy; beta-thalassaemia; cystic fibrosis; prematurity; blood loss;
KW   spinal cord injury; aging; neoplastic disease; erythropoiesis; EPO-R.
XX
OS   Synthetic.
XX
PN   WO9640749-A1.
XX
PD   19-DEC-1996.
XX
PF   07-JUN-1996;   96WO-US009810.
XX
PR   07-JUN-1995;   95US-00484631.
PR   07-JUN-1995;   95US-00484635.
XX
PA   (JOHJ ) JOHNSON & JOHNSON CORP.
PA   (AFFY-) AFFYMAX TECHNOLOGIES NV.
XX
PI   Wrighton NC,  Dower WJ,  Chang RS,  Kashyap AK,  Jolliffe LK;
PI   Johnson D,  Mulcahy L;
XX
DR   WPI; 1997-052225/05.
XX
PT   Erythropoietin receptor binding peptide - useful for treating disorders
PT   characterised by deficiency of EPO, or low or defective red blood cell
PT   population.
XX
PS   Disclosure; Page 25; 95pp; English.
XX
CC   The invention describes a peptide of 10-40 amino acid residues which
CC   binds to erythropoietin (EPO) receptor and which includes the amino acid
CC   sequence Cys-Xaa1-Xaa2-Gly-Pro-Xaa3-Thr-Trp-Xaa4-Cys, where Xaa1 = Arg,
CC   His, Leu or Trp, Xaa2 = Met, Phe or Ile, Xaa3 = 1 of the 20 genetically
CC   coded L-amino acids, and Xaa4 = Asp, Glu, Ile, Leu or Val. Optionally,
CC   the peptide may be cyclised or dimerised. The peptide can be used to
CC   treat a patient having a disorder characterised by a deficiency of EPO or
CC   a low or defective red blood cell population. It can be used to treat end
CC   stage renal failure or dialysis; anaemia associated with AIDS, autoimmune
CC   disease, chronic inflammatory diseases or malignancy; beta-thalassaemia;
CC   cystic fibrosis; early anaemia of prematurity; spinal cord injury; acute
CC   blood loss; aging; and neoplastic disease states accompanied by abnormal
CC   erythropoiesis. The peptides can also be used as reagents for detecting
CC   EPO receptors on living cells, in biological fluids, in tissue
CC   homogenates, etc. Sequences AAY26352-548 are representative peptides
CC   falling within the above peptide motif and isolated during the affinity
```

```
CC   selection process
XX
SQ   Sequence 20 AA;

  Query Match             93.6%;  Score 44;  DB 2;  Length 20;
  Best Local Similarity   60.0%;  Pred. No. 2.1;
  Matches    6;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 CXXGPXTWXC 10
              |  ||  || |
Db          7 CAFGPVTWEC 16


RESULT 11
AAY13696
ID   AAY13696 standard; peptide; 20 AA.
XX
AC   AAY13696;
XX
DT   06-SEP-1999  (first entry)
XX
DE   Erythropoietin receptor (EPO-R) binding peptide.
XX
KW   Erythropoietin; EPO; receptor; EPO deficiency; renal failure; AIDS;
KW   dialysis; anaemia; autoimmune disease; chronic inflammatory disease;
KW   malignancy; beta-thalassaemia; cystic fibrosis; prematurity; blood loss;
KW   spinal cord injury; aging; neoplastic disease; erythropoiesis; EPO-R.
XX
OS   Synthetic.
XX
PN   WO9640749-A1.
XX
PD   19-DEC-1996.
XX
PF   07-JUN-1996;   96WO-US009810.
XX
PR   07-JUN-1995;   95US-00484631.
PR   07-JUN-1995;   95US-00484635.
XX
PA   (JOHJ ) JOHNSON & JOHNSON CORP.
PA   (AFFY-) AFFYMAX TECHNOLOGIES NV.
XX
PI   Wrighton NC,  Dower WJ,  Chang RS,  Kashyap AK,  Jolliffe LK;
PI   Johnson D,  Mulcahy L;
XX
DR   WPI; 1997-052225/05.
XX
PT   Erythropoietin receptor binding peptide - useful for treating disorders
PT   characterised by deficiency of EPO, or low or defective red blood cell
PT   population.
XX
PS   Disclosure; Fig 2; 95pp; English.
XX
CC   The invention describes a peptide of 10-40 amino acid residues which
CC   binds to erythropoietin (EPO) receptor and which includes the amino acid
CC   sequence Cys-Xaa1-Xaa2-Gly-Pro-Xaa3-Thr-Trp-Xaa4-Cys, where Xaa1 = Arg,
CC   His, Leu or Trp, Xaa2 = Met, Phe or Ile, Xaa3 = 1 of the 20 genetically
CC   coded L-amino acids, and Xaa4 = Asp, Glu, Ile, Leu or Val. Optionally,
CC   the peptide may be cyclised or dimerised. The peptide can be used to
CC   treat a patient having a disorder characterised by a deficiency of EPO or
CC   a low or defective red blood cell population. It can be used to treat end
CC   stage renal failure or dialysis; anaemia associated with AIDS, autoimmune
CC   disease, chronic inflammatory diseases or malignancy; beta-thalassaemia;
CC   cystic fibrosis; early anaemia of prematurity; spinal cord injury; acute
CC   blood loss; aging; and neoplastic disease states accompanied by abnormal
CC   erythropoiesis. The peptides can also be used as reagents for detecting
CC   EPO receptors on living cells, in biological fluids, in tissue
CC   homogenates, etc. Sequences AAY13662-735 are representative peptides of
CC   the invention
XX
SQ   Sequence 20 AA;

  Query Match             93.6%;  Score 44;  DB 2;  Length 20;
  Best Local Similarity   60.0%;  Pred. No. 2.1;
  Matches    6;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 CXXGPXTWXC 10
              |  ||  || |
Db          6 CRMGPTTWEC 15


RESULT 12
AAY13650
ID   AAY13650 standard; peptide; 20 AA.
XX
AC   AAY13650;
XX
DT   06-SEP-1999  (first entry)
XX
DE   Erythropoietin receptor (EPO-R) binding peptide.
XX
KW   Erythropoietin; EPO; receptor; EPO deficiency; renal failure; AIDS;
KW   dialysis; anaemia; autoimmune disease; chronic inflammatory disease;
KW   malignancy; beta-thalassaemia; cystic fibrosis; prematurity; blood loss;
KW   spinal cord injury; aging; neoplastic disease; erythropoiesis; EPO-R.
XX
OS   Synthetic.
XX
PN   WO9640749-A1.
XX
PD   19-DEC-1996.
XX
PF   07-JUN-1996;   96WO-US009810.
XX
PR   07-JUN-1995;   95US-00484631.
PR   07-JUN-1995;   95US-00484635.
XX
PA   (JOHJ ) JOHNSON & JOHNSON CORP.
PA   (AFFY-) AFFYMAX TECHNOLOGIES NV.
XX
PI   Wrighton NC,  Dower WJ,  Chang RS,  Kashyap AK,  Jolliffe LK;
PI   Johnson D,  Mulcahy L;
XX
DR   WPI; 1997-052225/05.
XX
PT   Erythropoietin receptor binding peptide - useful for treating disorders
PT   characterised by deficiency of EPO, or low or defective red blood cell
PT   population.
XX
PS   Claim 6; Page 68; 95pp; English.
XX
CC   The invention describes a peptide of 10-40 amino acid residues which
CC   binds to erythropoietin (EPO) receptor and which includes the amino acid
CC   sequence Cys-Xaa1-Xaa2-Gly-Pro-Xaa3-Thr-Trp-Xaa4-Cys, where Xaa1 = Arg,
CC   His, Leu or Trp, Xaa2 = Met, Phe or Ile, Xaa3 = 1 of the 20 genetically
CC   coded L-amino acids, and Xaa4 = Asp, Glu, Ile, Leu or Val. Optionally,
CC   the peptide may be cyclised or dimerised. The peptide can be used to
CC   treat a patient having a disorder characterised by a deficiency of EPO or
CC   a low or defective red blood cell population. It can be used to treat end
CC   stage renal failure or dialysis; anaemia associated with AIDS, autoimmune
CC   disease, chronic inflammatory diseases or malignancy; beta-thalassaemia;
CC   cystic fibrosis; early anaemia of prematurity; spinal cord injury; acute
CC   blood loss; aging; and neoplastic disease states accompanied by abnormal
CC   erythropoiesis. The peptides can also be used as reagents for detecting
CC   EPO receptors on living cells, in biological fluids, in tissue
CC   homogenates, etc. Sequences AAY13624-661 represent specifc examples of
CC   EPO-R binding peptides
XX
SQ   Sequence 20 AA;

  Query Match             93.6%;  Score 44;  DB 2;  Length 20;
  Best Local Similarity   60.0%;  Pred. No. 2.1;
  Matches    6;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 CXXGPXTWXC 10
              |  ||  || |
```

```
Db          6 CHFGPATWVC 15


RESULT 13
AAY13728
ID   AAY13728 standard; peptide; 20 AA.
XX
AC   AAY13728;
XX
DT   06-SEP-1999  (first entry)
XX
DE   Erythropoietin receptor (EPO-R) binding peptide.
XX
KW   Erythropoietin; EPO; receptor; EPO deficiency; renal failure; AIDS;
KW   dialysis; anaemia; autoimmune disease; chronic inflammatory disease;
KW   malignancy; beta-thalassaemia; cystic fibrosis; prematurity; blood loss;
KW   spinal cord injury; aging; neoplastic disease; erythropoiesis; EPO-R.
XX
OS   Synthetic.
XX
PN   WO9640749-A1.
XX
PD   19-DEC-1996.
XX
PF   07-JUN-1996;   96WO-US009810.
XX
PR   07-JUN-1995;   95US-00484631.
PR   07-JUN-1995;   95US-00484635.
XX
PA   (JOHJ ) JOHNSON & JOHNSON CORP.
PA   (AFFY-) AFFYMAX TECHNOLOGIES NV.
XX
PI   Wrighton NC,  Dower WJ,  Chang RS,  Kashyap AK,  Jolliffe LK;
PI   Johnson D,  Mulcahy L;
XX
DR   WPI; 1997-052225/05.
XX
PT   Erythropoietin receptor binding peptide - useful for treating disorders
PT   characterised by deficiency of EPO, or low or defective red blood cell
PT   population.
XX
PS   Disclosure; Fig 2; 95pp; English.
XX
CC   The invention describes a peptide of 10-40 amino acid residues which
CC   binds to erythropoietin (EPO) receptor and which includes the amino acid
CC   sequence Cys-Xaa1-Xaa2-Gly-Pro-Xaa3-Thr-Trp-Xaa4-Cys, where Xaa1 = Arg,
CC   His, Leu or Trp, Xaa2 = Met, Phe or Ile, Xaa3 = 1 of the 20 genetically
CC   coded L-amino acids, and Xaa4 = Asp, Glu, Ile, Leu or Val. Optionally,
CC   the peptide may be cyclised or dimerised. The peptide can be used to
CC   treat a patient having a disorder characterised by a deficiency of EPO or
CC   a low or defective red blood cell population. It can be used to treat end
CC   stage renal failure or dialysis; anaemia associated with AIDS, autoimmune
CC   disease, chronic inflammatory diseases or malignancy; beta-thalassaemia;
CC   cystic fibrosis; early anaemia of prematurity; spinal cord injury; acute
CC   blood loss; aging; and neoplastic disease states accompanied by abnormal
CC   erythropoiesis. The peptides can also be used as reagents for detecting
CC   EPO receptors on living cells, in biological fluids, in tissue
CC   homogenates, etc. Sequences AAY13662-735 are representative peptides of
CC   the invention
XX
SQ   Sequence 20 AA;

  Query Match             93.6%;  Score 44;  DB 2;  Length 20;
  Best Local Similarity   60.0%;  Pred. No. 2.1;
  Matches    6;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 CXXGPXTWXC 10
              |  ||  ||  |
Db          6 CRMGPTTWVC 15


RESULT 14
AAY13688
ID   AAY13688 standard; peptide; 20 AA.
XX
AC   AAY13688;
XX
DT   06-SEP-1999  (first entry)
XX
DE   Erythropoietin receptor (EPO-R) binding peptide.
XX
KW   Erythropoietin; EPO; receptor; EPO deficiency; renal failure; AIDS;
KW   dialysis; anaemia; autoimmune disease; chronic inflammatory disease;
KW   malignancy; beta-thalassaemia; cystic fibrosis; prematurity; blood loss;
KW   spinal cord injury; aging; neoplastic disease; erythropoiesis; EPO-R.
XX
OS   Synthetic.
XX
PN   WO9640749-A1.
XX
PD   19-DEC-1996.
XX
PF   07-JUN-1996;   96WO-US009810.
XX
PR   07-JUN-1995;   95US-00484631.
PR   07-JUN-1995;   95US-00484635.
XX
PA   (JOHJ ) JOHNSON & JOHNSON CORP.
PA   (AFFY-) AFFYMAX TECHNOLOGIES NV.
XX
PI   Wrighton NC,  Dower WJ,  Chang RS,  Kashyap AK,  Jolliffe LK;
PI   Johnson D,  Mulcahy L;
XX
DR   WPI; 1997-052225/05.
XX
PT   Erythropoietin receptor binding peptide - useful for treating disorders
PT   characterised by deficiency of EPO, or low or defective red blood cell
PT   population.
XX
PS   Disclosure; Fig 2; 95pp; English.
XX
CC   The invention describes a peptide of 10-40 amino acid residues which
CC   binds to erythropoietin (EPO) receptor and which includes the amino acid
CC   sequence Cys-Xaa1-Xaa2-Gly-Pro-Xaa3-Thr-Trp-Xaa4-Cys, where Xaa1 = Arg,
CC   His, Leu or Trp, Xaa2 = Met, Phe or Ile, Xaa3 = 1 of the 20 genetically
CC   coded L-amino acids, and Xaa4 = Asp, Glu, Ile, Leu or Val. Optionally,
CC   the peptide may be cyclised or dimerised. The peptide can be used to
CC   treat a patient having a disorder characterised by a deficiency of EPO or
CC   a low or defective red blood cell population. It can be used to treat end
CC   stage renal failure or dialysis; anaemia associated with AIDS, autoimmune
CC   disease, chronic inflammatory diseases or malignancy; beta-thalassaemia;
CC   cystic fibrosis; early anaemia of prematurity; spinal cord injury; acute
CC   blood loss; aging; and neoplastic disease states accompanied by abnormal
CC   erythropoiesis. The peptides can also be used as reagents for detecting
CC   EPO receptors on living cells, in biological fluids, in tissue
CC   homogenates, etc. Sequences AAY13662-735 are representative peptides of
CC   the invention
XX
SQ   Sequence 20 AA;

  Query Match             93.6%;  Score 44;  DB 2;  Length 20;
  Best Local Similarity   60.0%;  Pred. No. 2.1;
  Matches    6;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 CXXGPXTWXC 10
              |  ||  ||  |
Db          6 CRMGPSTWEC 15


RESULT 15
AAY13687
ID   AAY13687 standard; peptide; 20 AA.
XX
AC   AAY13687;
XX
DT   06-SEP-1999  (first entry)
```

XX
DE Erythropoietin receptor (EPO-R) binding peptide.
XX
KW Erythropoietin; EPO; receptor; EPO deficiency; renal failure; AIDS;
KW dialysis; anaemia; autoimmune disease; chronic inflammatory disease;
KW malignancy; beta-thalassaemia; cystic fibrosis; prematurity; blood loss;
KW spinal cord injury; aging; neoplastic disease; erythropoiesis; EPO-R.
XX
OS Synthetic.
XX
PN WO9640749-A1.
XX
PD 19-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US009810.
XX
PR 07-JUN-1995; 95US-00484631.
PR 07-JUN-1995; 95US-00484635.
XX
PA (JOHJ) JOHNSON & JOHNSON CORP.
PA (AFFY-) AFFYMAX TECHNOLOGIES NV.
XX
PI Wrighton NC, Dower WJ, Chang RS, Kashyap AK, Jolliffe LK;
PI Johnson D, Mulcahy L;
XX
DR WPI; 1997-052225/05.
XX
PT Erythropoietin receptor binding peptide - useful for treating disorders
PT characterised by deficiency of EPO, or low or defective red blood cell
PT population.
XX
PS Disclosure; Fig 2; 95pp; English.
XX
CC The invention describes a peptide of 10-40 amino acid residues which
CC binds to erythropoietin (EPO) receptor and which includes the amino acid
CC sequence Cys-Xaa1-Xaa2-Gly-Pro-Xaa3-Thr-Trp-Xaa4-Cys, where Xaa1 = Arg,
CC His, Leu or Trp, Xaa2 = Met, Phe or Ile, Xaa3 = 1 of the 20 genetically
CC coded L-amino acids, and Xaa4 = Asp, Glu, Ile, Leu or Val. Optionally,
CC the peptide may be cyclised or dimerised. The peptide can be used to
CC treat a patient having a disorder characterised by a deficiency of EPO or
CC a low or defective red blood cell population. It can be used to treat end
CC stage renal failure or dialysis; anaemia associated with AIDS, autoimmune
CC disease, chronic inflammatory diseases or malignancy; beta-thalassaemia;
CC cystic fibrosis; early anaemia of prematurity; spinal cord injury; acute
CC blood loss; aging; and neoplastic disease states accompanied by abnormal
CC erythropoiesis. The peptides can also be used as reagents for detecting
CC EPO receptors on living cells, in biological fluids, in tissue
CC homogenates, etc. Sequences AAY13662-735 are representative peptides of
CC the invention
XX
SQ Sequence 20 AA;

Query Match 93.6%; Score 44; DB 2; Length 20;
Best Local Similarity 60.0%; Pred. No. 2.1;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy          1 CXXGPXTWXC 10
              |  ||  || |
Db          6 CRMGPTTWVC 15
```

Search completed: March 31, 2006, 16:22:26
Job time : 38.5572 secs

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: March 31, 2006, 16:22:51 ; Search time 6.21891 Seconds
(without alignments)
154.717 Million cell updates/sec

Title: US-10-609-217-421
Perfect score: 47
Sequence: 1 CXXGPXTWXC 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	45	95.7	19	1	EWSMAN	ancovenin - Strept
2	37	78.7	460	2	S06022	regulatory protein
3	37	78.7	475	2	H84137	hypothetical prote
4	36	76.6	123	2	I52427	guanine-nucleotide
5	36	76.6	123	2	S29714	guanine-nucleotide
6	35	74.5	318	2	E87929	protein T22H2.6 [i
7	35	74.5	345	2	T25138	hypothetical prote
8	35	74.5	358	2	T25137	hypothetical prote
9	35	74.5	2531	2	S18188	notch protein homo
10	35	74.5	2531	2	A46019	notch-1 protein -
11	35	74.5	2555	2	A40043	notch protein homo
12	34	72.3	19	1	EWSMCN	cinnamycin - Strep
13	34	72.3	78	1	EWSMYG	cinnamycin precurs
14	33	70.2	217	2	H86188	protein T25N20.5 [
15	33	70.2	266	2	H86407	F3H9.15 protein -
16	33	70.2	449	2	AC0234	probable exported
17	33	70.2	449	2	T47039	hypothetical prote
18	32.5	69.1	1661	2	T31330	head-activator bin
19	32.5	69.1	4543	1	A53102	alpha-2-macroglobu
20	32	68.1	279	2	G71429	hypothetical prote
21	32	68.1	292	2	S60997	ARG11 protein - ye
22	32	68.1	307	2	C81862	conserved hypothet
23	32	68.1	307	2	D81082	conserved hypothet
24	32	68.1	540	2	S72233	transcription fact
25	32	68.1	568	2	JC5629	mullerian-inhibiti
26	32	68.1	704	2	F86146	hypothetical prote
27	32	68.1	733	2	A97415	hypothetical prote
28	32	68.1	840	2	T02164	hypothetical prote
29	31.5	67.0	1149	2	I38006	M130 antigen precu
30	31.5	67.0	1151	2	I38004	M130 antigen precu
31	31.5	67.0	1156	2	I38005	M130 antigen precu
32	31.5	67.0	4753	1	A47437	LDL-receptor-relat
33	31	66.0	76	2	AH2120	hypothetical prote
34	31	66.0	126	2	S54062	hypothetical prote
35	31	66.0	180	2	T37095	hypothetical prote
36	31	66.0	211	2	C96539	hypothetical prote
37	31	66.0	290	2	E71256	probable P26 - syp
38	31	66.0	294	2	S13141	hypothetical prote
39	31	66.0	298	2	S12579	carbonate dehydrat
40	31	66.0	301	2	T27648	hypothetical prote
41	31	66.0	373	1	WMBET6	UL16 protein - hum
42	31	66.0	415	2	PC4407	envelope protein -
43	31	66.0	423	2	C86198	hypothetical prote
44	31	66.0	433	2	T39745	hypothetical prote
45	31	66.0	444	2	AD1823	hypothetical prote

ALIGNMENTS

RESULT 1
EWSMAN
ancovenin - Streptomyces sp. (strain A647P-2)
C;Species: Streptomyces sp.
C;Date: 12-May-1994 #sequence_revision 19-May-1994 #text_change 09-Jul-2004
C;Accession: A61284
R;Wakamiya, T.; Ueki, Y.; Shiba, T.; Kido, Y.; Motoki, Y.
Tetrahedron Lett. 26, 665-668, 1985
A;Title: The structure of ancovenin, a new peptide inhibitor of angiotensin I converting
A;Reference number: A61284
A;Accession: A61284
A;Molecule type: protein
A;Residues: 1-19 <WAK>
A;Cross-references: UNIPROT:P38655; UNIPARC:UPI0000052CC3
C;Superfamily: cinnamycin precursor
C;Keywords: antibiotic; lanthionine
F;1-18/Cross-link: (2S,3S,6R)-3-methyl-lanthionine (Cys-Thr) #status experimental
F;4-14/Cross-link: sn-(2S,6R)-lanthionine (Ser-Cys) #status experimental
F;5-11/Cross-link: (2S,3S,6R)-3-methyl-lanthionine (Cys-Thr) #status experimental
F;6/Modified site: dehydroalanine (Ser) #status experimental

Query Match 95.7%; Score 45; DB 1; Length 19;
Best Local Similarity 60.0%; Pred. No. 0.033;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy          1 CXXGPXTWXC 10
              |  || || |
Db          5 CSPGPLTWSC 14
```

RESULT 2
S06022
regulatory protein O2 - maize
C;Species: Zea mays (maize)
C;Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 31-Dec-2004
C;Accession: S06022; S06009
R;Hartings, H.; Maddaloni, M.; Lazzaroni, N.; di Fonzo, N.; Motto, M.; Salamini, F.; Thor
EMBO J. 8, 2795-2801, 1989
A;Title: The O2 gene which regulates zein deposition in maize endosperm encodes a protei
A;Reference number: S06022; MUID:90059860; PMID:2479535
A;Accession: S06022
A;Molecule type: mRNA
A;Residues: 1-460 <HAR>
A;Cross-references: UNIPROT:P12959; UNIPARC:UPI000016E05D; GB:X16618; NID:g22383; PIDN:C
R;Maddaloni, M.; di Fonzo, N.; Hartings, H.; Lazzaroni, N.; Salamini, F.; Thompson, R.;
Nucleic Acids Res. 17, 7532, 1989
A;Title: The sequence of the zein regulatory gene opaque-2 (O2) of Zea Mays.
A;Reference number: S06009; MUID:90016825; PMID:2798113
A;Accession: S06009
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-22,29-149,'D',151-460 <MAD>

```
A;Cross-references: UNIPARC:UPI00001794F4; EMBL:X15544
C;Genetics:
A;Gene: opaque 2
A;Map position: 7
A;Introns: 148/3; 168/3; 238/2; 263/3; 305/3
C;Superfamily: BZIP protein; fos/jun DNA-binding domain homology
C;Keywords: DNA binding; nucleus; transcription regulation
F;227-267/Domain: fos/jun DNA-binding domain homology <FJD>

  Query Match             78.7%;  Score 37;  DB 2;  Length 460;
  Best Local Similarity   71.4%;  Pred. No. 18;
  Matches    5;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 GPXTWXC 10
              || || |
Db        436 GPYTWTC 442


RESULT 3
H84137
hypothetical protein BH3904 [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C;Accession: H84137
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: H84137
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-475 <STO>
A;Cross-references: UNIPROT:Q9K628; UNIPARC:UPI00000C432F; GB:AP001520; GB:BA000004; NID
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH3904

  Query Match             78.7%;  Score 37;  DB 2;  Length 475;
  Best Local Similarity   62.5%;  Pred. No. 18;
  Matches    5;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 CXXGPXTW 8
              |  || ||
Db        156 CAGGPSTW 163


RESULT 4
I52427
guanine-nucleotide-releasing protein Mss4 - human
C;Species: Homo sapiens (man)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C;Accession: I52427
R;Yu, H.; Schreiber, S.L.
Biochemistry 34, 9103-9110, 1995
A;Title: Cloning, Zn2+ binding, and structural characterization of the guanine nucleotid
A;Reference number: I52427; MUID:95345082; PMID:7619808
A;Accession: I52427
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-123 <RES>
A;Cross-references: UNIPROT:P47224; UNIPARC:UPI00001117CC; GB:S78873; NID:g1037135; PIDN
C;Genetics:
A;Gene: GDB:MSS4
A;Cross-references: GDB:683578

  Query Match             76.6%;  Score 36;  DB 2;  Length 123;
  Best Local Similarity   50.0%;  Pred. No. 8.2;
  Matches    5;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 CXXGPXTWXC 10
              |  ||  | |
Db         97 CEIGPIGWHC 106


RESULT 5
S29714
guanine-nucleotide-releasing protein mss4 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C;Accession: S29714
R;Burton, J.; Roberts, D.; Montaldi, M.; Novick, P.; de Camilli, P.
Nature 361, 464-467, 1993
A;Title: A mammalian guanine-nucleotide-releasing protein enhances function of yeast sec
A;Reference number: S29714; MUID:93156814; PMID:8429887
A;Accession: S29714
A;Molecule type: mRNA
A;Residues: 1-123 <BUR>
A;Cross-references: UNIPROT:Q08326; UNIPARC:UPI000012F68D; EMBL:X70496; NID:g313871; PIDN
C;Genetics:
A;Gene: mss4

  Query Match             76.6%;  Score 36;  DB 2;  Length 123;
  Best Local Similarity   50.0%;  Pred. No. 8.2;
  Matches    5;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 CXXGPXTWXC 10
              |  ||  | |
Db         97 CEIGPIGWHC 106


RESULT 6
E87929
protein T22H2.6 [imported] - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Dec-2002
C;Accession: E87929
R;anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
A;Reference number: A75000; MUID:99069613; PMID:9851916
A;Note: see websites genome.wustl.edu/gsc/C_elegans/ and www_sanger.ac.uk/Projects/C_eleg
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A;Accession: E87929
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-318 <STO>
A;Cross-references: UNIPARC:UPI0000177C8F; GB:chr_I; PIDN:CAB04752.1; PID:g3880056; GSPDB
C;Genetics:
A;Gene: T22H2.6
A;Map position: 1
C;Superfamily: protein T22H2.6

  Query Match             74.5%;  Score 35;  DB 2;  Length 318;
  Best Local Similarity   50.0%;  Pred. No. 29;
  Matches    5;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 CXXGPXTWXC 10
              |  |  || |
Db         71 CKLGDNTWGC 80


RESULT 7
T25138
hypothetical protein T22H2.6b - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T25138
R;Lennard, N.
submitted to the EMBL Data Library, November 1996
A;Reference number: Z19985
A;Accession: T25138
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-345 <WIL>
A;Cross-references: UNIPROT:Q9U362; UNIPARC:UPI000002A1D2; EMBL:Z81595; PIDN:CAB54305.1;
A;Experimental source: clone T22H2
```

```
C;Genetics:
A;Gene: CESP:T22H2.6b
A;Map position: 1
A;Introns: 93/3; 232/3; 314/3
C;Superfamily: protein T22H2.6

  Query Match             74.5%;  Score 35;  DB 2;  Length 345;
  Best Local Similarity   50.0%;  Pred. No. 32;
  Matches    5;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 CXXGPXTWXC 10
              |  |  || |
Db        111 CKLGDNTWGC 120


RESULT 8
T25137
hypothetical protein T22H2.6a - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T25137
R;Lennard, N.
submitted to the EMBL Data Library, November 1996
A;Reference number: Z19985
A;Accession: T25137
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-358 <WIL>
A;Cross-references: UNIPROT:Q9U362; UNIPARC:UPI000008667D; EMBL:Z81595; PIDN:CAB54304.1;
A;Experimental source: clone T22H2
C;Genetics:
A;Gene: CESP:T22H2.6a
A;Map position: 1
A;Introns: 93/3; 232/3; 314/3
C;Superfamily: protein T22H2.6

  Query Match             74.5%;  Score 35;  DB 2;  Length 358;
  Best Local Similarity   50.0%;  Pred. No. 33;
  Matches    5;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 CXXGPXTWXC 10
              |  |  || |
Db        111 CKLGDNTWGC 120


RESULT 9
S18188
notch protein homolog - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 02-Aug-2002
C;Accession: S18188
R;Weinmaster, G.; Roberts, V.J.; Lemke, G.
Development 113, 199-205, 1991
A;Title: A homolog of Drosophila Notch expressed during mammalian development.
A;Reference number: S18188; MUID:92111383; PMID:1764995
A;Accession: S18188
A;Molecule type: mRNA
A;Residues: 1-2531 <WEI>
A;Cross-references: UNIPARC:UPI0000177456; EMBL:X57405; NID:g57634; PID:g57635
C;Superfamily: notch protein; ankyrin repeat homology; EGF homology
F;987-1018/Domain: EGF homology <EGF1>
F;1025-1056/Domain: EGF homology <EGF>
F;1233-1264/Domain: EGF homology <EGF2>
F;1917-1949/Domain: ankyrin repeat homology <AN1>
F;1950-1982/Domain: ankyrin repeat homology <AN2>
F;1984-2016/Domain: ankyrin repeat homology <AN3>
F;2017-2049/Domain: ankyrin repeat homology <AN4>
F;2050-2082/Domain: ankyrin repeat homology <AN5>

  Query Match             74.5%;  Score 35;  DB 2;  Length 2531;
  Best Local Similarity   50.0%;  Pred. No. 1.9e+02;
  Matches    5;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 CXXGPXTWXC 10
              |  || |: |
Db        543 CLDGPNTYTC 552


RESULT 10
A46019
notch-1 protein - mouse
N;Alternate names: motch protein
C;Species: Mus musculus (house mouse)
C;Date: 22-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 05-Oct-2004
C;Accession: A46019; S25144; C49175; B46438; A46438; PH1569; S32109
R;del Amo, F.F.; Gendron-Maguire, M.; Swiatek, P.J.; Jenkins, N.A.; Copeland, N.G.; Gridl
Genomics 15, 259-264, 1993
A;Title: Cloning, analysis, and chromosomal localization of Notch-1, a mouse homolog of D
A;Reference number: A46019; MUID:93194170; PMID:8449489
A;Accession: A46019
A;Status: not compared with conceptual translation
A;Molecule type: nucleic acid
A;Residues: 1-2531 <DEL>
A;Cross-references: UNIPROT:Q01705; UNIPARC:UPI000002922B; GB:Z11886; GB:S47228; NID:g288
A;Note: sequence extracted from NCBI backbone (NCBIP:127318)
R;Franco del Amo, F.; Smith, D.E.; Swiatek, P.J.; Gendron-Maguire, M.; Greenspan, R.J.; M
submitted to the EMBL Data Library, April 1992
A;Description: Expression pattern of Motch, a mouse homolog of Drosophila Notch, suggeste
A;Reference number: S25144
A;Accession: S25144
A;Molecule type: mRNA
A;Residues: 1551-2108,'Q',2110-2114,'ALP',2118-2170 <FRA>
A;Cross-references: UNIPARC:UPI0000177461; EMBL:Z11886
R;Lardelli, M.; Lendahl, U.
Exp. Cell Res. 204, 364-372, 1993
A;Title: Motch A and Motch B--two mouse Notch homologues coexpressed in a wide variety of
A;Reference number: A49175; MUID:93178563; PMID:8440332
A;Accession: C49175
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1161-1547 <LAR>
A;Cross-references: UNIPARC:UPI0000177462; EMBL:X68278; NID:g287987; PIDN:CAA48339.1; PID
A;Experimental source: embryo
A;Note: sequence extracted from NCBI backbone (NCBIP:126159)
R;Kopan, R.; Weintraub, H.
J. Cell Biol. 121, 631-641, 1993
A;Title: Mouse notch: expression in hair follicles correlates with cell fate determinatio
A;Reference number: A46438; MUID:93252998; PMID:8486742
A;Accession: B46438
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1865-1932,'RR',1935-1937,'L',1938-1967,'I',1969-2044,'IE',2047-2052,'S',2054
A;Cross-references: UNIPARC:UPI0000177463
A;Experimental source: embryo
A;Note: sequence extracted from NCBI backbone (NCBIN:131246, NCBIP:131247)
C;Comment: This protein has many EGF repeats and lin-12[#1172]/Notch repeats.
C;Comment: This protein is one of the neurogenic proteins controlling the decision betwee
C;Genetics:
A;Gene: notch-1
A;Map position: 2
A;Note: proximal region of chromosome 2
C;Superfamily: notch protein; ankyrin repeat homology; EGF homology
F;106-138/Domain: EGF homology <EGF1>
F;144-175/Domain: EGF homology <EG01>
F;222-254/Domain: EGF homology <EGF2>
F;261-292/Domain: EGF homology <EG02>
F;339-370/Domain: EGF homology <EG03>
F;416-449/Domain: EGF homology <EGF3>
F;456-487/Domain: EGF homology <EG04>
F;494-525/Domain: EGF homology <EG05>
F;532-563/Domain: EGF homology <EG06>
F;607-638/Domain: EGF homology <EG07>
F;682-713/Domain: EGF homology <EG08>
F;757-788/Domain: EGF homology <EG09>
F;795-826/Domain: EGF homology <EG10>
F;873-904/Domain: EGF homology <EG11>
```

F;911-942/Domain: EGF homology <EG12>
F;949-980/Domain: EGF homology <EG13>
F;987-1018/Domain: EGF homology <EG14>
F;1025-1056/Domain: EGF homology <EG15>
F;1063-1094/Domain: EGF homology <EG16>
F;1149-1180/Domain: EGF homology <EG17>
F;1187-1218/Domain: EGF homology <EG18>
F;1233-1264/Domain: EGF homology <EGF4>
F;1352-1383/Domain: EGF homology <EG19>
F;1391-1425/Domain: EGF homology <EGF>
F;1917-1948/Domain: ankyrin repeat homology <AN1>
F;1949-1981/Domain: ankyrin repeat homology <AN2>
F;1983-2015/Domain: ankyrin repeat homology <AN3>
F;2016-2048/Domain: ankyrin repeat homology <AN4>
F;2049-2081/Domain: ankyrin repeat homology <AN5>

Query Match 74.5%; Score 35; DB 2; Length 2531;
Best Local Similarity 50.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CXXGPXTWXC 10
Db 543 CLDGPNTYTC 552

RESULT 11
A40043
notch protein homolog TAN-1 precursor - human
C;Species: Homo sapiens (man)
C;Date: 21-Apr-1992 #sequence_revision 21-Apr-1992 #text_change 05-Oct-2004
C;Accession: A40043
R;Ellisen, L.W.; Bird, J.; West, D.C.; Soreng, A.L.; Reynolds, T.C.; Smith, S.D.; Sklar,
Cell 66, 649-661, 1991
A;Title: TAN-1, the human homolog of the Drosophila Notch gene, is broken by chromosomal
A;Reference number: A40043; MUID:91347367; PMID:1831692
A;Accession: A40043
A;Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
A;Molecule type: mRNA
A;Residues: 1-2555 <ELL>
A;Cross-references: UNIPARC:UPI0000177455; GB:M73980
C;Superfamily: notch protein; ankyrin repeat homology; EGF homology
F;261-292/Domain: EGF homology <EGX1>
F;494-525/Domain: EGF homology <EGF1>
F;987-1018/Domain: EGF homology <EGX2>
F;1149-1180/Domain: EGF homology <EGF>
F;1187-1218/Domain: EGF homology <EGF3>
F;1233-1264/Domain: EGF homology <EGX3>
F;1927-1959/Domain: ankyrin repeat homology <AN1>
F;1960-1992/Domain: ankyrin repeat homology <AN2>
F;1994-2026/Domain: ankyrin repeat homology <AN3>
F;2027-2059/Domain: ankyrin repeat homology <AN4>
F;2060-2092/Domain: ankyrin repeat homology <AN5>

Query Match 74.5%; Score 35; DB 2; Length 2555;
Best Local Similarity 50.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CXXGPXTWXC 10
Db 542 CLDGPNTYTC 551

RESULT 12
EWSMCN
cinnamycin - Streptoverticillium cinnamoneum
N;Alternate names: lanthiopeptin; lantibiotic Ro 09-0198
C;Species: Streptoverticillium cinnamoneum
C;Date: 30-Sep-1993 #sequence_revision 12-May-1994 #text_change 09-Jul-2004
C;Accession: A45767
R;Naruse, N.; Tenmyo, O.; Tomita, K.; Konishi, M.; Miyaki, T.; Kawaguchi, H.; Fukase, K.
J. Antibiot. 42, 837-845, 1989
A;Title: Lanthiopeptin, a new peptide antibiotic. Production, isolation and properties c
A;Reference number: A45767; MUID:89291558; PMID:2544544
A;Accession: A45767
A;Molecule type: protein
A;Residues: 1-19 <NAR>
A;Cross-references: UNIPROT:P29827; UNIPARC:UPI0000052CBF
R;Wakamiya, T.; Fukase, K.; Naruse, N.; Konishi, M.; Shiba, T.
Tetrahedron Lett. 29, 4771-4772, 1988
A;Title: Lanthiopeptin, a new peptide effective against Herpes simplex virus: structural
A;Reference number: A53359
A;Contents: annotation; strain L337-2
C;Superfamily: cinnamycin precursor
C;Keywords: antibiotic; beta-hydroxyaspartic acid; lanthionine
F;1-18/Cross-link: (2S,3S,6R)-3-methyl-lanthionine (Cys-Thr) #status experimental
F;4-14/Cross-link: sn-(2S,6R)-lanthionine (Ser-Cys) #status experimental
F;5-11/Cross-link: (2S,3S,6R)-3-methyl-lanthionine (Cys-Thr) #status experimental
F;6-19/Cross-link: (2Xi,9S)-lysinoalanine (Ser-Lys) #status experimental
F;15/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status experimental

Query Match 72.3%; Score 34; DB 1; Length 19;
Best Local Similarity 50.0%; Pred. No. 3.6;
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CXXGPXTWXC 10
Db 5 CSFGPFTFVC 14

RESULT 13
EWSMYG
cinnamycin precursor - Streptoverticillium griseoverticillatum
N;Alternate names: lanthiopeptin; lantibiotic Ro 09-0198
C;Species: Streptoverticillium griseoverticillatum
C;Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 09-Jul-2004
C;Accession: S17181; A60555
R;Kaletta, C.; Entian, K.D.; Jung, G.
Eur. J. Biochem. 199, 411-415, 1991
A;Title: Prepeptide sequence of cinnamycin (Ro 09-0198): the first structural gene of a c
A;Reference number: S17181; MUID:91301152; PMID:2070795
A;Accession: S17181
A;Molecule type: DNA
A;Residues: 1-78 <KAL>
A;Cross-references: UNIPROT:P29827; UNIPARC:UPI000000D539; EMBL:X58545; NID:g47089; PIDN:
R;Kessler, H.; Steuernagel, S.; Will, M.; Jung, G.; Kellner, R.; Gillessen, D.; Kamiyama,
Helv. Chim. Acta 71, 1924-1929, 1988
A;Title: The structure of the polycyclic nonadecapeptide Ro 09-0198.
A;Reference number: A60555
A;Accession: A60555
A;Molecule type: protein
A;Residues: 60-78 <KES>
A;Cross-references: UNIPARC:UPI0000052CBF
C;Genetics:
A;Gene: cinA; rocA
C;Superfamily: cinnamycin precursor
C;Keywords: antibiotic; beta-hydroxyaspartic acid; lanthionine
F;1-59/Domain: propeptide #status predicted <PRO>
F;60-78/Product: cinnamycin #status experimental <MAT>
F;60-77/Cross-link: (2S,3S,6R)-3-methyl-lanthionine (Cys-Thr) #status experimental
F;63-73/Cross-link: sn-(2S,6R)-lanthionine (Ser-Cys) #status experimental
F;64-70/Cross-link: (2S,3S,6R)-3-methyl-lanthionine (Cys-Thr) #status experimental
F;65-78/Cross-link: (2Xi,9S)-lysinoalanine (Ser-Lys) #status experimental
F;74/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status experimental

Query Match 72.3%; Score 34; DB 1; Length 78;
Best Local Similarity 50.0%; Pred. No. 13;
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CXXGPXTWXC 10
Db 64 CSFGPFTFVC 73

RESULT 14
H86188
protein T25N20.5 [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: H86188
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: H86188
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-217 <STO>
A;Cross-references: UNIPROT:Q9LR50; UNIPARC:UPI00000AB85D; GB:AE005172; NID:g8778714; PI
C;Genetics:
A;Gene: T25N20.5
A;Map position: 1
C;Superfamily: Arabidopsis hypothetical protein F9F13.130

Query Match 70.2%; Score 33; DB 2; Length 217;
Best Local Similarity 57.1%; Pred. No. 49;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy          4 GPXTWXC 10
              || :| |
Db        153 GPASWIC 159
```

RESULT 15
H86407
F3H9.15 protein - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: H86407
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: H86407
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-266 <STO>
A;Cross-references: UNIPROT:Q9FZ93; UNIPARC:UPI00000A5EE3; GB:AE005172; NID:g9795618; PI
C;Genetics:
A;Map position: 1

Query Match 70.2%; Score 33; DB 2; Length 266;
Best Local Similarity 57.1%; Pred. No. 59;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy          4 GPXTWXC 10
              || :| |
Db         19 GPSSWLC 25
```

Search completed: March 31, 2006, 16:37:17
Job time : 7.21891 secs

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: March 31, 2006, 16:09:36 ; Search time 37.4627 Seconds
(without alignments)
188.328 Million cell updates/sec

Title: US-10-609-217-421
Perfect score: 47
Sequence: 1 CXXGPXTWXC 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	45	95.7	19	1	DURC_STRGP	P36503 streptomyce
2	45	95.7	19	1	LANC_STRS6	P38655 streptomyce
3	40	85.1	167	2	Q6ZW93_HUMAN	Q6zw93 homo sapien
4	40	85.1	352	2	Q4IMN3_GIBZE	Q4imn3 gibberella
5	39	83.0	172	2	Q6ZT75_HUMAN	Q6zt75 homo sapien
6	39	83.0	172	2	Q6ZWC2_HUMAN	Q6zwc2 homo sapien
7	39	83.0	180	2	Q4I355_GIBZE	Q4i355 gibberella
8	38	80.9	414	2	Q4SAV9_TETNG	Q4sav9 tetraodon n
9	37	78.7	61	2	O70227_RAT	O70227 rattus norv
10	37	78.7	157	2	Q6NEH5_CORDI	Q6neh5 corynebacte
11	37	78.7	389	2	Q84UZ1_CHLRE	Q84uz1 chlamydomon
12	37	78.7	453	1	OP2_MAIZE	P12959 zea mays (m
13	37	78.7	475	2	Q9K628_BACHD	Q9k628 bacillus ha
14	37	78.7	556	2	Q84UZ4_CHLRE	Q84uz4 chlamydomon
15	37	78.7	741	2	Q4QE93_LEIMA	Q4qe93 leishmania
16	37	78.7	775	2	Q4H6M9_9DEIO	Q4h6m9 deinococcus
17	37	78.7	1192	2	Q7D3A2_AGRT5	Q7d3a2 agrobacteri
18	37	78.7	2022	2	Q61J27_CAEBR	Q61j27 caenorhabdi
19	36	76.6	123	1	MSS4_HUMAN	P47224 homo sapien
20	36	76.6	123	1	MSS4_MOUSE	Q91x96 mus musculu
21	36	76.6	123	1	MSS4_RAT	Q08326 rattus norv
22	36	76.6	123	2	Q53EV1_HUMAN	Q53ev1 homo sapien
23	36	76.6	132	2	Q6DGQ2_BRARE	Q6dgq2 brachydanio
24	36	76.6	202	2	Q5YYW8_NOCFA	Q5yyw8 nocardia fa
25	36	76.6	234	2	Q6IG67_DROME	Q6ig67 drosophila
26	36	76.6	532	2	Q8WWV6_HUMAN	Q8wwv6 homo sapien
27	36	76.6	534	2	Q96SA2_HUMAN	Q96sa2 homo sapien
28	36	76.6	544	2	Q4SD11_TETNG	Q4sd11 tetraodon n
29	36	76.6	577	2	Q5REH9_PONPY	Q5reh9 pongo pygma
30	36	76.6	589	2	Q5R770_PONPY	Q5r770 pongo pygma
31	36	76.6	887	2	Q726W3_DESVH	Q726w3 desulfovibr
32	36	76.6	1623	2	Q4SS52_TETNG	Q4ss52 tetraodon n
33	35	74.5	168	2	Q4P877_USTMA	Q4p877 ustilago ma
34	35	74.5	174	2	Q523Y1_MAGGR	Q523y1 magnaporthe
35	35	74.5	191	2	Q9DED0_CHICK	Q9ded0 gallus gall
36	35	74.5	238	2	Q7X189_9BACT	Q7x189 leptospiril
37	35	74.5	319	2	Q4WXG8_ASPFU	Q4wxg8 aspergillus
38	35	74.5	345	2	Q7JKP2_CAEEL	Q7jkp2 caenorhabdi
39	35	74.5	358	2	Q9U362_CAEEL	Q9u362 caenorhabdi
40	35	74.5	479	2	Q7XFS0_ORYSA	Q7xfs0 oryza sativ
41	35	74.5	479	2	Q8S776_ORYSA	Q8s776 oryza sativ
42	35	74.5	564	2	Q5AQY7_EMENI	Q5aqy7 aspergillus
43	35	74.5	2027	2	Q756Z8_ASHGO	Q756z8 ashbya goss
44	35	74.5	2067	2	Q59ED8_HUMAN	Q59ed8 homo sapien
45	35	74.5	2516	2	Q7TQ52_MOUSE	Q7tq52 mus musculu

ALIGNMENTS

```
RESULT 1
DURC_STRGP
ID   DURC_STRGP     STANDARD;      PRT;    19 AA.
AC   P36503;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Lantibiotic duramycin C.
OS   Streptomyces griseoluteus.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=29306;
RN   [1]
RP   PROTEIN SEQUENCE.
RC   STRAIN=R2107;
RX   MEDLINE=91107436; PubMed=2125590;
RA   Fredenhagen A., Fendrich G., Marki F., Marki W., Gruner J.,
RA   Raschdorf F., Peter H.H.;
RT   "Duramycins B and C, two new lanthionine containing antibiotics as
RT   inhibitors of phospholipase A2. Structural revision of duramycin and
RT   cinnamycin.";
RL   J. Antibiot. 43:1403-1412(1990).
RN   [2]
RP   STRUCTURE BY NMR.
RA   Zimmermann N., Freund S., Fredenhagen A., Jung G.;
RT   "Solution structure of the lantibiotics duramycin B and C.";
RL   (In) Schneider C.H., Eberles A.N. (eds.);
RL   Peptides 1992, pp.519-520, Escom Science Publishers, Leiden (1993).
RN   [3]
RP   STRUCTURE BY NMR.
RX   MEDLINE=93387292; PubMed=8375380;
RA   Zimmermann N., Freund S., Fredenhagen A., Jung G.;
RT   "Solution structures of the lantibiotics duramycin B and C.";
RL   Eur. J. Biochem. 216:419-428(1993).
CC   -!- FUNCTION: Acts as inhibitor of phospholipase A2.
CC   -!- PTM: Maturation of lantibiotics involves the enzymic conversion of
CC       Thr, and Ser into dehydrated AA and the formation of thioether
CC       bonds with cysteine or the formation of dialkylamine bonds with
CC       lysine. This is followed by membrane translocation and cleavage of
CC       the modified precursor.
CC   -!- SIMILARITY: Belongs to the type B lantibiotic family.
CC   ----------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ----------------------------------------------------------------------
KW   Antibiotic; Antimicrobial; Bacteriocin; Direct protein sequencing;
KW   Lantibiotic; Thioether bond.
FT   CROSSLNK      1      18       Beta-methyllanthionine (Cys-Thr).
FT   CROSSLNK      4      14       Lanthionine (Ser-Cys).
FT   CROSSLNK      5      11       Beta-methyllanthionine (Cys-Thr).
FT   CROSSLNK      6      19       Lysinoalanine (Ser-Lys).
```

```
SQ   SEQUENCE   19 AA;  2007 MW;  E2404ECE3F95286A CRC64;

  Query Match             95.7%;  Score 45;  DB 1;  Length 19;
  Best Local Similarity   60.0%;  Pred. No. 0.21;
  Matches    6;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 CXXGPXTWXC 10
              |  ||  || |
Db          5 CSYGPLTWSC 14


RESULT 2
LANC_STRS6
ID   LANC_STRS6     STANDARD;      PRT;    19 AA.
AC   P38655;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Lantibiotic ancovenin.
OS   Streptomyces sp. (strain A647P-2).
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=72591;
RN   [1]
RP   PROTEIN SEQUENCE.
RA   Wakamiya T., Ueki Y., Shiba T., Kido Y., Motoki Y.;
RT   "The structure of ancovenin, a new peptide inhibitor of angiotensin I
RT   converting enzyme.";
RL   Tetrahedron Lett. 26:665-668(1985).
CC   -!- FUNCTION: Acts as an inhibitor of angiotensin I converting enzyme.
CC   -!- PTM: Maturation of lantibiotics involves the enzymic conversion of
CC       Thr, and Ser into dehydrated AA and the formation of thioether
CC       bonds with cysteine or the formation of dialkylamine bonds with
CC       lysine. This is followed by membrane translocation and cleavage of
CC       the modified precursor.
CC   -!- SIMILARITY: Belongs to the type B lantibiotic family.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   PIR; A61284; EWSMAN.
KW   Antibiotic; Antimicrobial; Bacteriocin; Direct protein sequencing;
KW   Lantibiotic; Thioether bond.
FT   CROSSLNK      1     18       Beta-methyllanthionine (Cys-Thr).
FT   CROSSLNK      4     14       Lanthionine (Ser-Cys).
FT   CROSSLNK      5     11       Beta-methyllanthionine (Cys-Thr).
FT   CROSSLNK      6     19       Lysinoalanine (Ser-Lys).
SQ   SEQUENCE   19 AA;  2033 MW;  F434299E2736286A CRC64;

  Query Match             95.7%;  Score 45;  DB 1;  Length 19;
  Best Local Similarity   60.0%;  Pred. No. 0.21;
  Matches    6;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 CXXGPXTWXC 10
              |  ||  || |
Db          5 CSFGPLTWSC 14


RESULT 3
Q6ZW93_HUMAN
ID   Q6ZW93_HUMAN PRELIMINARY;      PRT;   167 AA.
AC   Q6ZW93;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Hypothetical protein FLJ41423.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Hippocampus;
RA   Kawakami B., Sugiyama A., Takemoto M., Sugiyama T., Irie R.,
RA   Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA   Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA   Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,
RA   Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A., Suzuki Y.,
RA   Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
RL   Submitted (JUL-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK123417; BAC85611.1; -; mRNA.
SQ   SEQUENCE   167 AA;  17960 MW;  266132D59393C276 CRC64;

  Query Match             85.1%;  Score 40;  DB 2;  Length 167;
  Best Local Similarity   50.0%;  Pred. No. 13;
  Matches    5;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 CXXGPXTWXC 10
              |  ||  | |
Db         83 CRQGPSVWSC 92


RESULT 4
Q4IMN3_GIBZE
ID   Q4IMN3_GIBZE PRELIMINARY;      PRT;   352 AA.
AC   Q4IMN3;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Hypothetical protein.
GN   ORFNames=FG01525.1;
OS   Gibberella zeae PH-1.
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC   Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
OX   NCBI_TaxID=229533;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=PH-1;
RA   Birren B., Nusbaum C., Abouelleil A., Allen N., Anderson S.,
RA   Arachchi H.M., Barna N., Bastien V., Bloom T., Boguslavkiy L.,
RA   Boukhgalter B., Butler J., Calvo S.E., Camarata J., Chang J.,
RA   Choepel Y., Collymore A., Cook A., Cooke P., Corum B., DeArellano K.,
RA   Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
RA   Erickson J., Faro S., Ferreira P., FitzGerald M., Gage D., Galagan J.,
RA   Gardyna S., Gnerre S., Graham L., Grand-Pierre N., Hafez N.,
RA   Hagopian D., Hagos B., Hall J., Horton L., Hulme W., Iliev I.,
RA   Jaffe D., Johnson R., Jones C., Kamal M., Kamat A., Karatas A.,
RA   Kells C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
RA   Ma L.-J., Mabbitt R., MacLean C., Macdonald P., Major J., Manning J.,
RA   Matthews C., Mauceli E., McCarthy M., Meldrim J., Meneus L.,
RA   Mihova T., Mlenga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
RA   Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neil D.,
RA   Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,
RA   Rachupka A., Ramasamy U., Raymond C., Retta R., Rise C., Rogov P.,
RA   Roman J., Schauer S., Schupback R., Seaman S., Severy P., Smirnov S.,
RA   Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,
RA   Talamas J., Tesfaye S., Theodore J., Topham K., Travers M.,
RA   Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,
RA   Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
RA   Lander E.;
RT   "Fusarium graminearum genome sequence.";
RL   Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; AACM01000077; EAA68151.1; -; Genomic_DNA.
KW   Hypothetical protein.
SQ   SEQUENCE   352 AA;  38308 MW;  670BA49FC645A7F8 CRC64;

  Query Match             85.1%;  Score 40;  DB 2;  Length 352;
  Best Local Similarity   50.0%;  Pred. No. 26;
  Matches    5;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;
```

```
Qy          1 CXXGPXTWXC 10
              |   | || | |
Db        184 CTSNPSTWRC 193
```

RESULT 5
Q6ZT75_HUMAN
ID Q6ZT75_HUMAN PRELIMINARY; PRT; 172 AA.
AC Q6ZT75;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein FLJ44897.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Amygdala;
RA Oshima A., Takahashi-Fujii A., Tanase T., Imose N., Takeuchi K.,
RA Arita M., Musashino K., Yuuki H., Hara H., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,
RA Murakawa K., Kanehori K., Sugiyama A., Kawakami B., Suzuki Y.,
RA Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AK126845; BAC86719.1; -; mRNA.
SQ SEQUENCE 172 AA; 18807 MW; DFD5579875B25559 CRC64;

```
  Query Match             83.0%;  Score 39;  DB 2;  Length 172;
  Best Local Similarity   50.0%;  Pred. No. 21;
  Matches    5;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 CXXGPXTWXC 10
              |  || :| |
Db          9 CLCGPESWTC 18
```

RESULT 6
Q6ZWC2_HUMAN
ID Q6ZWC2_HUMAN PRELIMINARY; PRT; 172 AA.
AC Q6ZWC2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein FLJ41341.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Brain;
RA Tanigami A., Fujiwara T., Shibahara T., Goto Y., Hirao M., Shimizu F.,
RA Wakebe H., Ono T., Hishigaki H., Watanabe T., Ozaki K., Sugiyama T.,
RA Irie R., Otsuki T., Sato H., Ota T., Wakamatsu A., Ishii S.,
RA Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
RA Kimura K., Yamashita H., Matsuo K., Nakamura Y., Sekine M.,
RA Kikuchi H., Kanda K., Wagatsuma M., Murakawa K., Kanehori K.,
RA Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B., Suzuki Y.,
RA Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AK123335; BAC85582.1; -; mRNA.
SQ SEQUENCE 172 AA; 18777 MW; C565579875A8FFF8 CRC64;

```
  Query Match             83.0%;  Score 39;  DB 2;  Length 172;
  Best Local Similarity   50.0%;  Pred. No. 21;
  Matches    5;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 CXXGPXTWXC 10
              |  || :| |
Db          9 CLCGPESWTC 18
```

RESULT 7
Q4I355_GIBZE
ID Q4I355_GIBZE PRELIMINARY; PRT; 180 AA.
AC Q4I355;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Predicted protein.
GN ORFNames=FG08353.1;
OS Gibberella zeae PH-1.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
OX NCBI_TaxID=229533;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PH-1;
RA Birren B., Nusbaum C., Abouelleil A., Allen N., Anderson S.,
RA Arachchi H.M., Barna N., Bastien V., Bloom T., Boguslavkiy L.,
RA Boukhgalter B., Butler J., Calvo S.E., Camarata J., Chang J.,
RA Choepel Y., Collymore A., Cook A., Cooke P., Corum B., DeArellano K.,
RA Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
RA Erickson J., Faro S., Ferreira P., FitzGerald M., Gage D., Galagan J.,
RA Gardyna S., Gnerre S., Graham L., Grand-Pierre N., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Horton L., Hulme W., Iliev I.,
RA Jaffe D., Johnson R., Jones C., Kamal M., Kamat A., Karatas A.,
RA Kells C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
RA Ma L.-J., Mabbitt R., MacLean C., Macdonald P., Major J., Manning J.,
RA Matthews C., Mauceli E., McCarthy M., Meldrim J., Meneus L.,
RA Mihova T., Mlenga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
RA Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neil D.,
RA Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,
RA Rachupka A., Ramasamy U., Raymond C., Retta R., Rise C., Rogov P.,
RA Roman J., Schauer S., Schupback R., Seaman S., Severy P., Smirnov S.,
RA Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,
RA Talamas J., Tesfaye S., Theodore J., Topham K., Travers M.,
RA Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,
RA Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
RA Lander E.;
RT "Fusarium graminearum genome sequence.";
RL Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC   EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC   preliminary data.
DR EMBL; AACM01000335; EAA72141.1; -; Genomic_DNA.
SQ SEQUENCE 180 AA; 20463 MW; 94C7B5242FEE6ED9 CRC64;

```
  Query Match             83.0%;  Score 39;  DB 2;  Length 180;
  Best Local Similarity   50.0%;  Pred. No. 22;
  Matches    5;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 CXXGPXTWXC 10
              |  ||  | |
Db        101 CSFGPAPWEC 110
```

RESULT 8
Q4SAV9_TETNG
ID Q4SAV9_TETNG PRELIMINARY; PRT; 414 AA.
AC Q4SAV9;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome 3 SCAF14679, whole genome shotgun sequence.
GN ORFNames=GSTENG00021242001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

```
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC   Tetradontoidea; Tetraodontidae; Tetraodon.
OX   NCBI_TaxID=99883;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA   Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA   Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA   Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA   Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA   Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA   Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA   Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA   Kellis M., Volff JN., Guigo R., Zody M.C., Mesirov J.,
RA   Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA   Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA   Wincker P., Lander E.S., Weissenbach J., Roest Crollius H.;
RT   "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT   the early vertebrate proto-karyotype.";
RL   Nature 431:946-957(2004).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RG   Genoscope; Whitehead Institute Centre for Genome Research;
RL   Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; CAAE01014679; CAG02223.1; -; Genomic_DNA.
SQ   SEQUENCE   414 AA;  45368 MW;  0522D03EA381377E CRC64;

  Query Match             80.9%;  Score 38;  DB 2;  Length 414;
  Best Local Similarity   50.0%;  Pred. No. 71;
  Matches    5;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 CXXGPXTWXC 10
              |   | || |
Db        155 CRMSPSTWGC 164


RESULT 9
O70227_RAT
ID   O70227_RAT  PRELIMINARY;      PRT;    61 AA.
AC   O70227;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-AUG-1998 (TrEMBLrel. 07, Last annotation update)
DE   MARRLC9A (Fragment).
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC   Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Brain;
RA   Liao B.S., Jin W.L., Ju G.;
RL   Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF010444; AAC14892.1; -; mRNA.
FT   NON_TER       1      1
FT   NON_TER      61     61
SQ   SEQUENCE   61 AA;  6655 MW;  C8AF3B9CB8656126 CRC64;

  Query Match             78.7%;  Score 37;  DB 2;  Length 61;
  Best Local Similarity   50.0%;  Pred. No. 19;
  Matches    5;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 CXXGPXTWXC 10
              |  ||  | |
Db         23 CPRGPGQWAC 32


RESULT 10
Q6NEH5_CORDI
ID   Q6NEH5_CORDI PRELIMINARY;      PRT;   157 AA.
AC   Q6NEH5;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Putative integral membrane protein.
GN   OrderedLocusNames=DIP2299;
OS   Corynebacterium diphtheriae.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX   NCBI_TaxID=1717;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Biotype gravis / NCTC 13129;
RX   MEDLINE=22965443; PubMed=14602910; DOI=10.1093/nar/gkg874;
RA   Cerdeno-Tarraga A.-M., Efstratiou A., Dover L.G., Holden M.T.G.,
RA   Pallen M.J., Bentley S.D., Besra G.S., Churcher C.M., James K.D.,
RA   De Zoysa A., Chillingworth T., Cronin A., Dowd L., Feltwell T.,
RA   Hamlin N., Holroyd S., Jagels K., Moule S., Quail M.A.,
RA   Rabbinowitsch E., Rutherford K.M., Thomson N.R., Unwin L.,
RA   Whitehead S., Barrell B.G., Parkhill J.;
RT   "The complete genome sequence and analysis of Corynebacterium
RT   diphtheriae NCTC13129.";
RL   Nucleic Acids Res. 31:6516-6523(2003).
DR   EMBL; BX248360; CAE50822.1; -; Genomic_DNA.
KW   Complete proteome.
SQ   SEQUENCE   157 AA;  17842 MW;  6B28DD518B7D5CD5 CRC64;

  Query Match             78.7%;  Score 37;  DB 2;  Length 157;
  Best Local Similarity   50.0%;  Pred. No. 45;
  Matches    5;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 CXXGPXTWXC 10
              |  |  || |
Db         82 CESGDATWIC 91


RESULT 11
Q84UZ1_CHLRE
ID   Q84UZ1_CHLRE PRELIMINARY;      PRT;   389 AA.
AC   Q84UZ1;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Putative gag protein.
OS   Chlamydomonas reinhardtii.
OC   Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae;
OC   Chlamydomonadales; Chlamydomonadaceae; Chlamydomonas.
OX   NCBI_TaxID=3055;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Perez-Alegre M., Fernandez E.;
RL   Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY227353; AAO73553.1; -; Genomic_DNA.
SQ   SEQUENCE   389 AA;  44595 MW;  7B9C6FA3BA70D2DA CRC64;

  Query Match             78.7%;  Score 37;  DB 2;  Length 389;
  Best Local Similarity   71.4%;  Pred. No. 1e+02;
  Matches    5;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 GPXTWXC 10
              || || |
Db        337 GPRTWTC 343


RESULT 12
OP2_MAIZE
ID   OP2_MAIZE      STANDARD;      PRT;   453 AA.
AC   P12959;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-JAN-1990 (Rel. 13, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Opaque-2 regulatory protein.
```

```
GN   Name=O2;
OS   Zea mays (Maize).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX   NCBI_TaxID=4577;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=AC 1503 GM 1407;
RX   MEDLINE=90016825; PubMed=2798113;
RA   Maddaloni M., di Fonzo N., Hartings H., Lazzaroni N., Salamini F.,
RA   Thompson R.D., Motto M.;
RT   "The sequence of the zein regulatory gene opaque-2 (O2) of Zea mays.";
RL   Nucleic Acids Res. 17:7532-7532(1989).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=AC 1503 GM 1407; TISSUE=Seed endosperm;
RX   MEDLINE=90059860; PubMed=2479535;
RA   Hartings H., Maddaloni M., Lazzaroni N., di Fonzo N., Motto M.,
RA   Salamini F., Thompson R.D.;
RT   "The O2 gene which regulates zein deposition in maize endosperm
RT   encodes a protein with structural homologies to transcriptional
RT   activators.";
RL   EMBO J. 8:2795-2801(1989).
RN   [3]
RP   FUNCTION.
RX   MEDLINE=91160516; PubMed=2001677;
RA   Lohmer S., Maddaloni M., Motto M., di Fonzo N., Hartings H.,
RA   Salamini F., Thompson R.D.;
RT   "The maize regulatory locus Opaque-2 encodes a DNA-binding protein
RT   which activates the transcription of the b-32 gene.";
RL   EMBO J. 10:617-624(1991).
RN   [4]
RP   TISSUE SPECIFICITY, AND INTERACTION WITH PBF.
RX   MEDLINE=97352860; PubMed=9207153; DOI=10.1073/pnas.94.14.7685;
RA   Vicente-Carbajosa J., Moose S.P., Parsons R.L., Schmidt R.J.;
RT   "A maize zinc-finger protein binds the prolamin box in zein gene
RT   promoters and interacts with the basic leucine zipper transcriptional
RT   activator Opaque-2.";
RL   Proc. Natl. Acad. Sci. U.S.A. 94:7685-7690(1997).
CC   -!- FUNCTION: Involved in the regulation of the endosperm-specific
CC       production of albumin b-32 and other zein proteins. It is a trans-
CC       acting transcriptional activator that binds to the consensus
CC       sequence 5'-GATGAYRTGR-3'.
CC   -!- SUBUNIT: Interacts with the Dof zinc finger protein PBF.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- TISSUE SPECIFICITY: Seed endosperm.
CC   -!- SIMILARITY: Belongs to the bZIP family.
CC   -!- SIMILARITY: Contains 1 bZIP domain.
CC   ----------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ----------------------------------------------------------------------
DR   EMBL; X15544; CAA33550.1; -; Genomic_DNA.
DR   EMBL; X16618; CAA34614.1; -; mRNA.
DR   PIR; S06022; S06022.
DR   HSSP; P03069; 2DGC.
DR   TRANSFAC; T00668; -.
DR   Gramene; P12959; -.
DR   MaizeDB; 24976; -.
DR   InterPro; IPR011616; bZIP_1.
DR   InterPro; IPR004827; TF_bZIP.
DR   Pfam; PF00170; bZIP_1; 1.
DR   SMART; SM00338; BRLZ; 1.
DR   PROSITE; PS50217; BZIP; 1.
DR   PROSITE; PS00036; BZIP_BASIC; 1.
KW   Activator; DNA-binding; Nuclear protein; Transcription;
KW   Transcription regulation.
FT   DOMAIN      253    274       Leucine-zipper.
FT   DNA_BIND    228    246       Basic motif.
FT   CONFLICT     26     26       E -> EPEPEPE (in Ref. 2).
FT   CONFLICT    144    144       D -> A (in Ref. 2).
FT   CONFLICT    231    231       K -> KR (in Ref. 2).
SQ   SEQUENCE   453 AA;  49357 MW;  513A8AB8D5ABD999 CRC64;

  Query Match             78.7%;  Score 37;  DB 1;  Length 453;
  Best Local Similarity   71.4%;  Pred. No. 1.2e+02;
  Matches    5;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 GPXTWXC 10
              || || |
Db        429 GPYTWTC 435


RESULT 13
Q9K628_BACHD
ID   Q9K628_BACHD PRELIMINARY;      PRT;   475 AA.
AC   Q9K628;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   BH3904 protein.
GN   OrderedLocusNames=BH3904;
OS   Bacillus halodurans.
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=86665;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=C-125 / JCM 9153;
RX   MEDLINE=20512582; PubMed=11058132; DOI=10.1093/nar/28.21.4317;
RA   Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA   Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA   Horikoshi K.;
RT   "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT   halodurans and genomic sequence comparison with Bacillus subtilis.";
RL   Nucleic Acids Res. 28:4317-4331(2000).
DR   EMBL; BA000004; BAB07623.1; -; Genomic_DNA.
DR   PIR; H84137; H84137.
DR   InterPro; IPR008557; DUF839_bac.
DR   InterPro; IPR006311; Tat.
DR   Pfam; PF05787; DUF839; 1.
DR   TIGRFAMs; TIGR01409; TAT_signal_seq; 1.
KW   Complete proteome.
SQ   SEQUENCE   475 AA;  52013 MW;  8FBACFD6185533F3 CRC64;

  Query Match             78.7%;  Score 37;  DB 2;  Length 475;
  Best Local Similarity   62.5%;  Pred. No. 1.2e+02;
  Matches    5;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 CXXGPXTW 8
              |  || ||
Db        156 CAGGPSTW 163


RESULT 14
Q84UZ4_CHLRE
ID   Q84UZ4_CHLRE PRELIMINARY;      PRT;   556 AA.
AC   Q84UZ4;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Putative gag-polyprotein.
OS   Chlamydomonas reinhardtii.
OC   Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae;
OC   Chlamydomonadales; Chlamydomonadaceae; Chlamydomonas.
OX   NCBI_TaxID=3055;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Perez-Alegre M., Fernandez E.;
RL   Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY227352; AAO73550.1; -; Genomic_DNA.
KW   Polyprotein.
SQ   SEQUENCE   556 AA;  62299 MW;  52B80C1EB66C58D2 CRC64;
```

Query Match 78.7%; Score 37; DB 2; Length 556;
Best Local Similarity 71.4%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          4 GPXTWXC 10
              || || |
Db        504 GPRTWTC 510
```

RESULT 15
Q4QE93_LEIMA
ID Q4QE93_LEIMA PRELIMINARY; PRT; 741 AA.
AC Q4QE93;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=LmjF17.0980;
OS Leishmania major.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Friedlin;
RA Peacock C.S., Murphy L., Ivens A.C., Berriman M., Blackwell J.,
RA Smith D., Collins M., Fosker N., Harris D., Oliver K., O'Neil S.,
RA Saunders D., Seeger K., Warren T., Rajandream M., and Barrell B.G.;
RL Submitted (JUN-2005) to the EMBL/GenBank/DDBJ databases.
DR EMBL; CT005256; CAJ03767.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 741 AA; 82904 MW; A6167BC4712A5E01 CRC64;

Query Match 78.7%; Score 37; DB 2; Length 741;
Best Local Similarity 50.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy          1 CXXGPXTWXC 10
              |  |  || |
Db         25 CLKGESTWSC 34
```

Search completed: March 31, 2006, 16:35:09
Job time : 38.4627 secs